Development of an image processing method for automated, non-invasive and scale-independent monitoring of adherent cell cultures

Nicolas Jaccard

Thesis submitted for the degree of Doctor of Philosophy

Department of Biochemical Engineering, University College London

Centre for Mathematics and Physics in the Life Sciences and Experimental Biology, University College London

25 February 15

Disclaimer

I, Nicolas Jaccard, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed NICOLAS JACCARD

Abstract

Adherent cell culture is a key experimental method for biological investigations in diverse areas such as developmental biology, drug discovery and biotechnology. Light microscopy-based methods, for example phase contrast microscopy (PCM), are routinely used for visual inspection of adherent cells cultured in transparent polymeric vessels. However, the outcome of such inspections is qualitative and highly subjective. Analytical methods that produce quantitative results can be used but often at the expense of culture integrity or viability.

In this work, an imaging-based strategy to adherent cell cultures monitoring was investigated. Automated image processing and analysis of PCM images enabled quantitative measurements of key cell culture characteristics. Two types of segmentation algorithms for the detection of cellular objects on PCM images were evaluated. The first one, based on contrast filters and dynamic programming was quick (<1s per 1280×960 image) and performed well for different cell lines, over a wide range of imaging conditions. The second approach, termed 'trainable segmentation', was based on machine learning using a variety of image features such as local structures and symmetries. It accommodated complex segmentation tasks while maintaining low processing times (<5s per 1280×960 image).

Based on the output from these segmentation algorithms, imaging-based monitoring of a large palette of cell responses was demonstrated, including proliferation, growth arrest, differentiation, and cell death. This approach is non-invasive and applicable to any transparent culture vessel, including microfabricated culture devices where a lack of suitable analytical methods often limits their applicability. This work was a significant contribution towards the establishment of robust, standardised, and affordable monitoring methods for adherent cell cultures. Finally, automated image processing was combined with computer-controlled cultures in small-scale devices. This provided a first demonstration of how adaptive culture protocols could be established; i.e. culture protocols which are based on cellular response instead of arbitrary time points.

Abstract word count: 299

Table of Contents

Disclaimer 2

Abstract 3

Table of Contents 4

List of Figures 7

List of Tables 11

List of Abbreviations and Acronyms 13

List of Symbols 14

Acknowledgments 15

1. Introduction and background 16
1.1 Adherent cell culture: current approaches and limitations 16
1.1.1 Typical adherent cell culture workflow and opportunities for quantification 17
1.1.2 Type of adherent cells and their applications 18
1.1.3 Environments for optimal cell growth and function 20
1.1.4 Culture systems for anchorage-dependent cells 22
1.1.5 A quantitative adherent cell culture framework 27
1.2 Non-invasive monitoring of live adherent cell cultures 28
1.2.1 Methods requiring detachment of the cells 29
1.2.2 Bright field and phase contrast methods 31
1.2.3 Digital holographic microscopy 36
1.2.4 Fluorescence microscopy 36
1.2.5 Raman microspectroscopy 41
1.2.6 Impedance sensing 42
1.2.7 Indirect monitoring 43
1.3 Automated processing and analysis of light microscopy images 44
1.3.1 Applications in adherent cell culture monitoring 44
1.3.2 Phase contrast microscopy image segmentation 46
1.3.3 Object detection and tracking on phase contrast microscopy images 55
1.3.4 Considerations for real-world applications 57
1.4 Towards a high performance and scale-independent monitoring methods for live adherent cells 58

2. Segmentation of phase contrast images using contrast filters 61
2.1 Description of PCM segmentation approaches based on contrast filters 62
2.1.1 Image pre-processing 63
2.1.2 Segmentation 63
2.1.3 Post-processing hole filling and small objects removal 68
2.1.4 *Post-hoc* halo correction 68
2.2 Methods for dataset generation and segmentation performance evaluation 70
2.2.1 Cell culture 70
2.2.2 Image acquisition 71
2.2.3 Datasets for performance evaluation 72

2.2.4 Evaluation of segmentation performance 73
2.2.5 Optimisation of segmentation parameters 77
2.3 Results of segmentation performance evaluation 80
2.3.1 Segmentation using contrast filters 80
2.3.2 Improvement of segmentation performance using post-hoc halo correction 84
2.3.3 Comparison with other methods 90
2.3.4 Generalisation to other cell lines and imaging conditions 91
2.4 Open-source implementation: PHANTAST 96
2.4.1 PHANTAST GUI for adherent culture PCM image processing 97
2.4.2 A tool for automated parameter estimation based on sparse annotations 98
2.4.3 ImageJ/Fiji plugin 101
2.5 Conclusion and outlook 102

3. Imaging-based monitoring of adherent cell culture 105
3.1 Deriving informative data from PCM image segmentation output 105
3.1.1 Confluency estimation 106
3.1.2 Morphometric analysis 112
3.1.3 Augmented fluorescence images 113
3.2 Application examples 114
3.2.1 Impact of culture medium exchange schedule on mESC growth kinetics 114
3.2.2 Transient cell response induced by environmental shock 117
3.2.3 Cytotoxicity assay for chemical stressor 121
3.2.4 Early and long-term differentiation monitoring 123
3.3 Towards a method for estimation of cell density from PCM images 129
3.4 Conclusion 135

4. Trainable segmentation of phase contrast microscopy images 136
4.1 Machine learning methods applied to image processing 138
4.1.1 Schemes for pixel-wise image segmentation using machine learning techniques 139
4.1.2 Pixel class label prediction using random decision forests 142
4.2 Computational methods 144
4.2.1 Features computation 144
4.2.2 Local feature histograms 146
4.2.3 Computation time for local image feature histograms construction 150
4.2.4 Random forest classifier 151
4.2.5 Segmentation process 151
4.2.6 Evaluation of segmentation performance 151
4.3 Results and discussion 152
4.3.1 Efficient computation of local image feature histograms 153
4.3.2 Trainable segmentation of mouse embryonic stem cells PCM images 157
4.3.3 Trainable segmentation of Human embryonic stem cell PCM images 168
4.4 Comparison with other trainable segmentation schemes 176
4.5 Conclusion and outlook 178

5. Integration of imaging-based monitoring with a microfabricated device for cell culture process development 184
5.1 Microfabricated device for cell culture process development 185
5.1.1 The analytical bottleneck and the imaging-based solution 186
5.2 Material and methods 187
5.2.1 Human embryonic stem cell cultures in the microfabricated device 187

5.2.2 Mouse embryonic stem cell cultures in the microfabricated device 188
5.2.3 LabVIEW routine for automated stage control and image acquisition 189
5.2.4 Image processing and analysis 190
5.3 Results and discussion 191
5.3.1 Intermittent monitoring of human embryonic stem cell cultures 191
5.3.2 Continuous monitoring of mouse embryonic stem cell cultures 194
5.4 Conclusion and outlook 202

6. Conclusion and recommendations 204
6.1 Two approaches to PCM images segmentation 204
6.2 Towards a standard monitoring approach for adherent cultures 207
6.3 The next step in adherent cell bioprocessing 209
6.4 Summary of main contributions, novelty, and advances 211

Glossary 214

Bibliography 216

Annexes 229

A. Digital microscopy images 229

B. Methods for the determination of the intensity gradient of a digital image 229
B.1.1 Using the finite difference method 230
B.1.2 Using convolution 230
B.1.3 Using convolution with a Gaussian kernel 232

C. Design and operation of the optimisation framework 233
C.1 Flowcharts describing framework operation 237
C.1.1 XML parsing and loading 237

D. Code listings 238
D.1 Simple Gaussian filter implementation 238

E. Standard deviation computation 239

List of Figures

Figure 1-1 A typical adherent cell culture workflow. 17
Figure 1-2 Reimagining of the Waddington's landscape to reflect modern understanding of cell plasticity. 19
Figure 1-3 Examples of different systems for adherent cell culture. 22
Figure 1-4 Examples of T-flasks used for routine maintenance of adherent cells. 23
Figure 1-5 Examples of well-plates used for parallelised cultures of adherent cells. 24
Figure 1-6 Data quality and disruption for a few adherent cell culture analytical techniques. 28
Figure 1-7 Schematic of a flow cytometer. 30
Figure 1-8 Light interaction with specimen. 32
Figure 1-9 Working principle of phase contrast microscopy (PCM). 33
Figure 1-10 Comparison of bright field and phase contrast microscopy. 33
Figure 1-11 Typical intensity profile in the neighbourhood of a cell on a PCM image. 34
Figure 1-12 Three different phase contrast microscopy setups 35
Figure 1-13 Wide-field fluorescence microscopy principles. 37
Figure 1-14 Strategies for fluorescence microscopy 39
Figure 1-15 Example of a PCM image and the corresponding binary segmentation output. 46
Figure 1-16 Examples of contrast filter approaches for PCM image segmentation. 48
Figure 1-17 Active contour models for PCM image segmentation. 50
Figure 1-18 Example of application of watershed transforms to image segmentation. 52
Figure 2-1 Visual features of mouse embryonic stem cells (mESCs) at various stages of a low seeding density expansion experiment. 62
Figure 2-2 Intensity histogram of a PCM image of a mES cell shortly after seeding. 64
Figure 2-3 Simple intensity threshold for PCM image segmentation. 64
Figure 2-4 Hard-edged local standard deviation (heSTDEV) filter output for a PCM image of a mES cell shortly after seeding. 66
Figure 2-5 Normalised soft-edged local standard deviation (nseSTDEV) filter output for a PCM image of a mES cell shortly after seeding 67
Figure 2-6 Range filter output for a PCM image of a mES cell shortly after seeding. 68
Figure 2-7 Method for *post-hoc* correction of halo artefacts. 69
Figure 2-8 Mouse embryonic stem cell (mESC) PCM image dataset. 73
Figure 2-9 Synthetic example illustrating the behaviour of the various ROCs considered. 75
Figure 2-10 Detailed analysis of the F-score behaviour using artificial examples. 76
Figure 2-11 Further investigation of the F-score behaviour. 77
Figure 2-12 Parameter tuning of the segmentation algorithm. 78
Figure 2-13 Assessment of segmentation performance for three contrast thresholding algorithms and one based on raw intensity thresholding only. 82
Figure 2-14 Examples of segmentation outputs for the various algorithms tested. 83
Figure 2-15 Illustration of the segmentation process using contrast filters and *post-hoc* halo correction. 84

Figure 2-16 Assessment of segmentation performance with halo removal for three contrast thresholding algorithms and one based on raw intensity thresholding only. 86
Figure 2-17 Examples of segmentation outputs for the various algorithms tested with halo removal. 88
Figure 2-18 Stability of segmentation performance for non-optimal intensity thresholding parameters. 89
Figure 2-19 Processing time of the iterative halo correction algorithm. 90
Figure 2-20 Comparison of the proposed segmentation approaches to methods previously described in the literature. 91
Figure 2-21 Segmentation of Chinese hamster ovary (CHO) cells and Human neuroblastoma (NB) PCM images. 92
Figure 2-22 Segmentation results for PCM images of a wide range of cell types. 93
Figure 2-23 Effect of the imaging setup on segmentation outcome. 94
Figure 2-24 Tolerance of the segmentation algorithm to variations in imaging conditions. 95
Figure 2-25 Screenshots of the PHANTAST graphical user interface (GUI). 97
Figure 2-26 Screenshot of the optimisation tool. 99
Figure 2-27 Parameter value optimisation based on sparse image annotation. 100
Figure 2-28 ImageJ/Fiji implementation of PHANTAST. 102
Figure 3-1 Evaluation of confluency determination variability. 107
Figure 3-2 Image confluency determination performance. 109
Figure 3-3 Example of image quality differences between PCM images acquired with and without condensation present on the culture vessel lid. 110
Figure 3-4 Evaluation of the impact of sampling error on culture confluency estimation. 112
Figure 3-5 Examples of morphological properties that can be extracted from the segmentation output. 112
Figure 3-6 Augmented fluorescence image (AFI) constructed by combining PCM image segmentation and fluorescence microscopy. 114
Figure 3-7 Experimental setup for the monitoring of the impact of medium exchange on the growth of mESCs. 115
Figure 3-8 Growth kinetics for various medium exchange regimes. 116
Figure 3-9 Experimental setup for the monitoring of cell response to environmental shock. .. 118
Figure 3-10 Transient cell response monitoring. 119
Figure 3-11 Further characterisation of the response of cells to environmental shock. 120
Figure 3-12 Experimental setup for imaging-based cytotoxicity assay. 121
Figure 3-13 Monitoring cell response to a chemical stressor. 122
Figure 3-14 Experimental setup for differentiation monitoring experiment. 123
Figure 3-15 Boxplots for the morphometric analysis of early differentiation events. 125
Figure 3-16 Time-course study of mESC differentiation events. 126
Figure 3-17 Augmented fluorescence images (AFIs) for the visualisation of spatial differentiation patterns. 127
Figure 3-18 PCM images of Oct-4 GiP cells after 280h of culture in expansion medium, spontaneous differentiation medium, and directed differentiation medium. 128
Figure 3-19 Non-linear relationship between image confluency and number of nuclei. 130
Figure 3-20 Packing-corrected image confluency. 131
Figure 3-21 The seven Basic Image Features (BIFs) and the corresponding colour code. 132

Figure 3-22 Basic Image Features (BIFs) as an alternative to fluorescent markers..... 132
Figure 3-23 Packing-corrected confluency for mESC culture cell density estimation. 133
Figure 4-1 Examples of structured background noise..... 136
Figure 4-2 A single cell line can exhibit a wide range of visual features. 137
Figure 4-3 PCM images of co-cultures are challenging to segment using conventional approaches..... 137
Figure 4-4 Data representation for classification of handwritten digits..... 139
Figure 4-5 Pixel-wise segmentation of an image using only pixel intensity as feature. 139
Figure 4-6 Pixel-wise segmentation of an image using multiple features for pixel classification. 140
Figure 4-7 Different feature scales reveal different information. 140
Figure 4-8 Pixel-wise segmentation of an image using multiple features each computed at two scales for pixel classification. 141
Figure 4-9 Pixel-wise segmentation of an image using multi-scale local image feature histograms. 141
Figure 4-10 Decision tree for classification. 142
Figure 4-11 Training and prediction phases of a random forest classifier for pixel classification (background vs foreground). 144
Figure 4-12 Bank of Derivative-of-Gaussian (DtG) filters used for the computation of BIFs. 145
Figure 4-13 The trainable segmentation scheme employed either discrete or interval histograms. 147
Figure 4-14 Image set used for the evaluation of Human embryonic stem cell (hESC) segmentation performance. 152
Figure 4-15 Proposed trainable segmentation scheme. 153
Figure 4-16 Impact of image size (A) and number of bins (B) on histogram computation time. 155
Figure 4-17 Impact of local neighbourhood size on histogram computation time for (A) 7 bins, (B) 28 bins and (C) 56 bins histograms. 156
Figure 4-18 Examples of features used for trainable segmentation of mESC PCM images. ... 158
Figure 4-19 Trainable segmentation of mESC PCM images based on intensity features. 160
Figure 4-20 Trainable segmentation of mESC PCM images based on contrast filter features. 162
Figure 4-21 Features importance for segmentation using contrast features..... 163
Figure 4-22 Trainable segmentation of mESC PCM images based on Basic Image Features (BIFs). 165
Figure 4-23 Trainable segmentation using combinations of features. 166
Figure 4-24 Feature importance when combining raw intensity, contrast and Basic Image Features at five different scales..... 166
Figure 4-25 Comparison of mESCs PCM images segmentation outputs between contrast thresholding and trainable segmentation..... 167
Figure 4-26 Segmentation of hESC PCM images..... 169
Figure 4-27 Examples of features used for trainable segmentation of hESC PCM images..... 169
Figure 4-28 Trainable segmentation of hESC PCM images based on pixel intensity values. .. 171
Figure 4-29 Trainable segmentation of hESC PCM images based on contrast features..... 173
Figure 4-30 Trainable segmentation of hESC PCM images based on Basic Image Features (BIFs). 175

Figure 4-31 Comparison of different trainable segmentation schemes for four full-resolution mESC images. 177
Figure 4-32 Comparison of different trainable segmentation schemes for three full-resolution hESC images. 178
Figure 4-33 Robustness to structured background noise and unusual cell visual features. 180
Figure 4-34 Trainable segmentation, given suitable annotations, outperforms contrast thresholding in extreme cases. 181
Figure 4-35 Bespoke graphical user interface for interactive trainable segmentation in MATLAB 182
Figure 5-1 The microfabricated culture device. 185
Figure 5-2 Setup for the intermittent monitoring of cultures of Human embryonic stem cells (hESC) in the microfabricated device. 191
Figure 5-3 Typical images of hESC co-cultured with MEFs in the microfabricated device (a-c) and a conventional culture dish (d-f). 192
Figure 5-4 Tracking dynamic culture events. 193
Figure 5-5 Approach for online monitoring of cultures of mouse embryonic stem cells (mESC) in the microfabricated device. 195
Figure 5-6 Integration of dissolved oxygen monitoring. 195
Figure 5-7 Whole culture chamber images at various time points during mESC cultures in the microfabricated culture device. 196
Figure 5-8 On-chip mESC culture confluency monitoring. 198
Figure 5-9 Determination of local growth patterns. 199
Figure 5-10 Finite element model of oxygen gradient in the microfabricated device. 201
Figure 6-1 Two fundamentally different approaches to PCM image segmentation. 204
Figure 6-2 Schematic detailing the design and operation of the optimisation framework 233
Figure 6-3 Relationship between the main classes of the optimisation framework 234
Figure 6-4 Framework GUI for parameter values optimisation 236

List of Tables

Table 1-1 Non-exhaustive summary of analytical methods for adherent cell culture characterisation 29
Table 1-2 Selection of publications describing methods for the segmentation of cells on phase contrast microscopy images. 47
Table 2-1 Embryonic stem cell maintenance medium formulation 71
Table 2-2 CHO and Neuroblastoma cell culture medium formulation 71
Table 2-3 Various imaging setups used during the course of this project 72
Table 2-4 Confusion matrix for the comparison of segmentation outputs with ground truth images 74
Table 2-5 Receiver operator characteristics (ROCs) for the analysis of the confusion matrix 74
Table 2-6 Explored parameter space for segmentation algorithms evaluation (without halo correction). 81
Table 2-7 Optimal processing parameters for PCM image segmentation using contrast filters. L 82
Table 2-8 Explored parameter space for segmentation algorithm (with halo correction) 85
Table 2-9 Optimal processing parameters for PCM image segmentation using contrast filters and *post-hoc* halo correction. 87
Table 3-1 Confluency estimation performance based on segmentation output from a normalised soft-edged local standard deviation algorithm with and without halo correction 109
Table 3-2 Number of images required at a 10× magnification to cover the whole culture area of commonly used culture vessels. 111
Table 3-3 Validation of the calibration process for culture cell density estimation based on packing-corrected confluency (PCC). 134
Table 4-1 Classification of BIFs according to which term results in the largest numerical value. The colour codes are used for graphical representation of BIFs image. 146
Table 4-2 Results for trainable segmentation of mESC PCM images based on raw pixel intensity values 159
Table 4-3 Results for trainable segmentation of mESC PCM images based on contrast features 163
Table 4-4 Results for trainable segmentation of mESC PCM images using Basic Image Features 164
Table 4-5 Best performing settings for each feature types considered for trainable segmentation of mESC PCM images 167
Table 4-6 Results for trainable segmentation of hESC PCM images based on raw intensity features 170
Table 4-7 Results for trainable segmentation of hESC PCM images based on contrast features 172
Table 4-8 Trainable segmentation of hESC PCM images based on basic image features 174
Table 4-9 Best performing settings for each feature types considered for hESC trainable segmentation 176

Table 4-10 Comparison of the proposed trainable segmentation scheme (multi-scale BIFs histograms) with schemes implemented in two other trainable segmentation software packages. 176
Table 5-1 Mouse embryonic fibroblast feeder cells maintenance medium formulation..... 188
Table 5-2 Human embryonic stem cell maintenance medium formulation 188
Table 5-3 Monitoring of hESC cultures in the microfabricated bioreactor and a conventional cell culture dish (single-well dish with a 2.89 cm^2 culture area) 193
Table 6-1 Comparison of contrast thresholding and trainable segmentation. 205
Table 6-2 Return types for framework operations 235
Table 6-3 Type of attributes for the specification of operation parameters 235

List of Abbreviations and Acronyms

CHO	Chinese hamster ovary (cells)
GUI	Graphical User Interface
DT(s)	Decision tree(s)
ECM	Extracellular matrix
hESC(s)	Human embryonic stem cells
heSTDEV	Hard-edged local standard deviation filter
heLH	Hard-edged local image feature histogram
LOOCV	Leave-one-out cross-validation
mESC(s)	Mouse embryonic stem cells
NB	Neuroblastoma (cells)
nseSTDEV	Normalised soft-edged local standard deviation filter
OOP	Object-oriented programming
PCM	Phase contrast microscopy
PSC(s)	Pluripotent stem cells
PDMS	poly(dimethylsiloxane)
PNG	Portable network graphics
RANGE	Range contrast filter
RF	Random forest
RGB	Red Green Blue (colour channels)
RMS	Raman microspectroscopy
ROCs	Receiver operating characteristics
seLH	Soft-edged local image feature histogram
STA	Statistics of local image feature histogram
TIF(F)	Tagged image file format
TTFR	Time to first result
VI	Virtual instrument
XML	Extensible markup language

List of Symbols

A_{FOH}	Hole filling area threshold
A_{RSO}	Small object removal area threshold
$b^{(k)}$	Binary mask used during histogram construction by convolution for the k-th bin
c_{ij}	Response to the convolution of an image by a Derivative of Gaussian kernel of order i and j in the x and y dimensions, respectively
$c^{(k)}$	Image resulting from convolution of $b^{(k)}$ with a Gaussian kernel for histogram construction using convolution
ε	Threshold parameter (applied to both intensity thresholding and Basic Image Features computation)
e	Segmentation error
$E[X]$	Expected value of a random variable X
$H(x,y)$	Local feature histogram for pixel located at coordinates x and y
G_σ	Gaussian kernel of standard deviation σ
G_σ^i	Derivative of Gaussian kernel of standard deviation σ and order i
$G(x,y)$	Binary image resulting from intensity thresholding of raw PCM image
$P_{culture}$	Precision of the confluency estimation for an entire culture (taking into account the uncertainty introduced by random sampling)
$P_{intrinsic}$	Intrinsic precision of the confluency estimation for a single image (**not** taking into account uncertainty introduced by random sampling)
$P_{sampling}$	Uncurtaining in confluency estimation introduced by random sampling across the culture area
σ	Standard deviation, also used to denote the scale of a feature
R_{MAX}	Fraction of object allowed to shrink during halo correction
$S(x,y)$	Standard deviation image
V	Illumination light voltage
V_{intra}	Intra-variability between human experts
V_{inter}	Inter-variability between human experts

Acknowledgments

I would like to thank both my supervisors, Dr. Nicolas Szita and Dr. Farlan Veraitch, for the time they put aside to provide invaluable insights. I would also like to acknowledge Dr. Lewis Griffin, who helped me tremendously with the image processing aspects of my work. Both Brian O'Sullivan and Ludmila Ruban provided me with their technical expertise during the course of my project.

I am also very grateful to the member of the Microfluidics group in the department of Biochemical Engineering in UCL for the fruitful discussions and general support: Alexandre Super, James Lawrence, Marcel Reichen, and Rhys Macown.

I would like to give special thanks to all those who have helped shape my scientific mind over the years. Dr Sebastien Chenuet in particular, whose limitless enthusiasm for science was contagious.

Thank you to my friends from Switzerland who supported me in my decision to undertake a PhD and believed in me: Benoit Pugin, Delphine Balmer, Fabien Rawyler, Ivan Küttel, Jonathan Ochs, and Olivier Feyer. I would also like to thank my friends and colleagues from CoMPLEX: David Sheehan, Donal O'Donoghue, Mirna Kovacevic, and everyone from the 2009 cohort.

Finally, and most importantly, thank you to my family, Bernadette, François, Axel, Fiona, and Abbygail, without whose love and support I would not have been able to carry on for the last four years.

Financial support from CoMPLEX, the EPSRC, and the British Heart Foundation is gratefully acknowledged.

Chapter 1

Introduction and background

1.1 Adherent cell culture: current approaches and limitations

Cell culture refers to the growth of disaggregated cells under controlled conditions outside of their natural environment (Langdon, 2004). It is a collection of techniques and apparatuses that enable the handling of cells in aseptic conditions. It represented a major step forward for experimental cell biology when it was first established in the early 20th century and underlined the reductionism approach, whereby individual cell populations were studied *in-vitro* (i.e. in a dish) in an effort to gain knowledge about the tissue or organ they were originally derived from (Carrel, 1912; Gallagher, 1999).

The term adherent cell culture is used in opposition to suspension cell culture, which as the name indicates refers to the growth of cells as free-floating particles. Suspension cultures are mostly limited to biotechnology and bioprocess applications where microbial or suspension-adapted animal cells are grown in continuously agitated culture vessels whose scale ranges from the millilitre to the thousands of litres (Doran, 2013). In contrast, adherent cell culture involves the growth of cells that are attached to a substrate. At the exception of some hematopoietic lineages, most somatic animal cells are anchorage-dependent and require attachment to a substrate to maintain viability and function (Jinno et al., 2001). Preventing cell from attaching to a growth surface was shown to induce apoptosis through a process mediated by trans-membrane receptors (Ruoslahti and Reed, 1994).

Adherent cell culture is an experimental framework for biology with many facets. In addition to cell handling and the culture vessels to perform cell culture itself, it also encompasses concepts such as analytical techniques, design of experiment, and data analysis. From the early days of painstaking microscopic observations (Schierbeek et al., 1959) to the '-omics' era and the big data paradigms (Berger et al., 2013), all aspects of cell culture saw incremental improvements and refinements to match the ever increasing complexity of the biological processes studied. However, current experimental approaches have limitations that can potentially hinder progress in certain areas of biology. In the past two decades, there have been concerted efforts (e.g. systems biology) to overcome these limitations by complementing traditional experimental

approaches with methods borrowed from other fields such as engineering, applied mathematics, material science, and computer science (Kitano, 2002; Kohl et al., 2010).

This work explores how such an inter-disciplinary approach could be taken to establish robust and systematic methods for the monitoring of adherent cell cultures. In order to provide a context for this research and its underlying motivations, key concepts of adherent cell cultures are first introduced, with a brief summary of current approaches and their limitations.

1.1.1 Typical adherent cell culture workflow and opportunities for quantification

Culture monitoring and characterisation is a key aspect of any adherent cell culture workflow, be it for decision-making or evaluation of experimental outcomes (Figure 1-1). Cell culture experiments are initiated by seeding the culture vessel with a single cell suspension (i.e. liquid suspension containing cells that are assumed to be separated from one another rather than in clumps). In order to ensure consistency between experiments, the number of cells used for seeding is determined using standard enumeration methods. This single cell suspension can also be used to establish baseline measurements for other characteristics that might be of interest. After seeding, the vessel is placed in an incubator. It is good practice to intermittently monitor the culture by visually inspecting the cells using a light microscope. At the end of the experiment, the cells are enzymatically dissociated and re-suspended. At this point, the single cell suspension can be used for further characterisation.

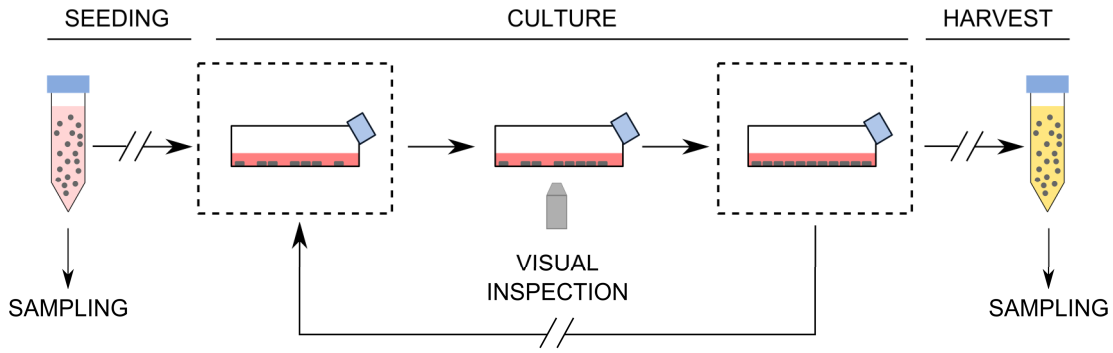


Figure 1-1 A typical adherent cell culture workflow. Cells are seeded into the culture vessel, which is then placed into an incubator (as represented by the dashed line). The culture is intermittently inspected using light microscopy approaches. At the end of the culture, the cells are harvested by enzymatic dissociation.

This typical workflow contains two steps where quantitative measurements are carried out: the start and the end of the culture. These measurements are performed on cells that have been detached prior to analysis and therefore cannot inform on key adherent cell characteristics such as spatial distribution or cellular object morphology. Subjective and qualitative evaluations of the culture by visual inspection often constitute the full extent of the data acquired during the course of an experiment. In many cases, there is thus little to no information on the changes that occur between the start and end points of a culture.

A number of analytical methods have been developed to complement visual inspection and provide quantitative measurements of adherent cell culture characteristics. Those will be discussed in detail in section 1.2. Intuitively, microscopy imaging data appears to be a reasonable starting point for the development of such a method as images are already being generated in most cases, but possibly not exploited to their fullest potential (this statement will be substantiated later in the text). This research project, together with other initiatives (see section 1.3), aims at taking advantage of this treasure chest of unused microscopy data to devise a method that enable rapid and non-invasive measurement of cell culture characteristics.

1.1.2 Type of adherent cells and their applications

There are three types of cells that are routinely used in adherent cell culture: primary cells, immortalized cell lines, and pluripotent stem cells. Each type of cell has its own set of advantages and limitations. This highlights the breadth of potential applications for the monitoring approach described in this work, as it can in principle be used with all three types of cells.

1.1.2.1 Primary cell lines

Cells that are put in culture directly after isolation from *in-vivo* material are known as primary cells (Alberts et al., 2002). These cells are the gold standard in many studies as their phenotype closely matches that of their *in-vivo* counterpart. When maintained past the initial primary culture, the cell population is then referred to as a cell line (Langdon, 2004). Cells tend to undergo drastic phenotypic drift as the culture progresses (Darling and Athanasiou, 2005; Yao et al., 2006). For many applications, it is therefore necessary to establish new primary cultures for each experiment to carry out, which is highly time consuming and leads to increased reliance on animal models. With the exception of cells derived from tumours, primary cell lines are usually finite and will stop proliferating within 50 generations, precluding long-term culture experiments (Evan and d'Adda di Fagagna, 2009; Hayflick, 1965).

1.1.2.2 Immortalised cell lines

Finite cell lines can be made continuous (or immortal) through a process known as transformation (or immortalisation), during which cells are treated with chemical carcinogens or viruses to effectively create a cell line that possess the characteristics of cancer cells (Smith et al., 1992). Immortalised cell lines greatly facilitate cell culture experiments and enable some degree of standardisation across laboratories. However, it is widely recognized that the immortalisation process drastically impacts on the phenotype of the cells and the resulting cell line will thus not necessarily retain the characteristics of the initial primary culture (Stampfer et al., 1997). Immortalized cell lines are extensively used for drug discovery and toxicity studies

(Allen et al., 2005). They also can be used as cell factories, cellular hosts engineered for optimal production of high value molecules such as vaccines (Genzel et al., 2014).

1.1.2.3 Pluripotent stem cells

The third category of cells is pluripotent stem cells (PSCs), a special case of primary cell lines. PSCs are isolated from the blastocyst of a developing embryo and are self-renewable, meaning they can divide and proliferate as long as suitable signals are present in their environment (Thomson et al., 1998). Moreover, PSCs can give rise to any cell type of an adult organism through a process known as differentiation (Odorico et al., 2001). Cells that are cultured *in-vitro* and present characteristics of PSCs are often referred to as embryonic stem cells (ESCs).

The notion that a cell's fate is sealed after it reached a terminally differentiated state (Figure 1-2) was challenged by Takahashi and colleagues when they successfully reprogrammed adult fibroblasts into pluripotent stem cells that closely resemble ESCs (Takahashi et al., 2007; Takahashi and Yamanaka, 2006). These induced pluripotent stem cells (iPSCs) alleviate the concerns over ethical issues and controversies regarding the origin of ESCs. In addition, the fact that any adult cell could potentially be reprogrammed and generate pluripotent cells paves the way for patient-specific therapies with reduced risks of immune response upon transplantation.

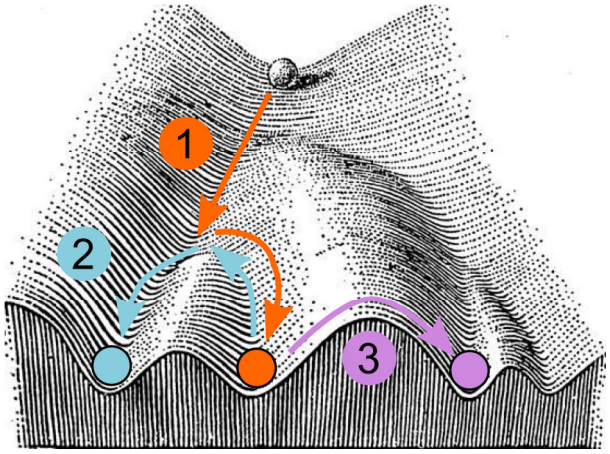


Figure 1-2 Reimagining of the Waddington's landscape to reflect modern understanding of cell plasticity. Differentiation (1), reprogramming and re-differentiation (2) and transdifferentiation (3). Original landscape drawing adapted from (Waddington, 1957).

The combination of self-renewal and differentiation potential makes PSCs a potentially unlimited pool of cells for developmental biology (Wobus and Boheler, 2005), drug discovery (Pouton and Haynes, 2007; Sartipy et al., 2007) and toxicity studies (Davila et al., 2004; Sison-Young et al., 2012). PSCs also have great potential for regenerative medicine approaches, whereby transplantation of PSC-derived progenitors or terminally differentiated cells could cure conditions that are not treatable by conventional means. Potential target illnesses include cancer (Mimeault et al., 2007), neurodegenerative disease (Goldman, 2005), HIV (Deeks and McCune, 2010) or cardiovascular disease (Segers and Lee, 2008).

The experimental component of this work was carried out mainly with mouse and human embryonic stem cells (mESCs and hESCs, respectively). This choice was motivated by their high relevance across many biological fields, though generalisation to other types of cells was also critical.

1.1.3 Environments for optimal cell growth and function

Cell culture relies on the ability to re-create environments *ex-vivo* that can sustain optimal growth and function of cells. The ability to do so appropriately will depend on the capabilities of the culture vessel and techniques employed, as further discussed in section 1.1.4. The type of environment required and complexity thereof will also be dictated by the type of cells being studied and the desired outcome. The cell microenvironment can be characterised by its physical, biochemical, and physicochemical properties (Young and Beebe, 2010; Zhang and van Noort, 2011).

1.1.3.1 Biochemical environment

The biochemical environment is usually defined by the composition of the culture medium in which cells are submerged and by molecules released by the cells. Basic culture medium formulation includes carbon sources (e.g. glucose), amino acids, vitamins and other nutrients. A major component of the culture media for adherent cells is serum, a complex mixture containing hormones, growth factors, and molecules that are required for proper cell attachment to the growth substrate (Gasser et al., 1985). Serum is costly, prone to batch-to-batch variability, and its animal origins as well as the uncertainties regarding its composition might be an issue, especially in a clinical context. Efforts were made to formulate serum-free media that contain the molecules usually present in serum but at defined concentrations. These defined media formulations have been widely available for suspension cultures but the move towards animal-free defined media remains a challenge for adherent cell culture, most likely due to the additional requirements of molecules that promote attachment (Petiot et al., 2010).

While the culture medium formulation is usually similar across most cell culture applications, it is often necessary to supplement it with a cocktail of molecules that are specific for the type of cell considered or the investigation to carry out. An example is the culture of mouse embryonic stem cells (mESCs), which requires the addition of leukaemia inhibitor factor (LIF) to the culture medium in order to promote the maintenance of pluripotency (Hayashi et al., 2007). In some situations, it is not possible to devise the cocktail of molecules that promote a certain cell function. A possible approach then consists in co-culturing the cell line being investigated with other cell types that are known to enhance said function *in-vivo* (in this case, the other cell population is usually referred to as feeders). For example, this approach can be used for the

long-term maintenance of hESC pluripotency in culture (Richards et al., 2002) or their differentiation (Bin et al., 2006).

The biochemical environment is very dynamic and changes continuously during the course of a culture experiment as nutrients get depleted and potentially toxic metabolic wastes accumulate. In most cases, its control is limited to bulk medium exchanges. Those can either be arbitrarily timed (e.g. once a day) or based on cues, most often a change in colour of a pH-sensitive dye such as phenol red. Neither is sufficient to maintain long-term optimal culture conditions, which require the use of more advanced strategies. One such strategy is the use of microfluidic culture devices and will be discussed in section 1.1.4.3.

1.1.3.2 Physicochemical environment

In the context of cell culture, physicochemical properties usually refer to temperature, pH, oxygen, and carbon dioxide (Young and Beebe, 2010; Zhang and van Noort, 2011). These are either directly or indirectly controlled by placing the culture vessel in an incubator. Temperature is usually kept constant at a physiological 37 °C and the gas atmosphere at ~20% O_2 and 5% CO_2. It is important to note that these values correspond to the atmosphere in the incubator and do not necessarily match what the cells actually experience (Cochran et al., 2006). In fact during most cell culture experiments, physicochemical properties of the cell microenvironment are not measured (with the exception of pH through the aforementioned colour indicator) and thus should not be considered as controlled culture parameters. Culture pH is usually controlled by a combination of the 5% CO_2 gaseous atmosphere and a bicarbonate buffer in the culture medium.

1.1.3.3 Mechanical environment

The physical environment of the cells encompasses both the extracellular matrix (ECM) and external mechanical forces (e.g. fluid induced shear stress, stretching). *In-vivo*, the composition of the ECM varies greatly depending on the tissue considered but is generally made up of proteoglycans, collagen, elastin, fibronectin or laminin (Alberts et al., 2002). Transmembrane integrins mediate the attachment and other interactions of the cells with the surrounding ECM (Leung-Hagesteijn et al., 1994). The composition of the ECM, as well as its physical properties (e.g. elasticity), drastically impact cell phenotype and function (Baharvand et al., 2005; Hirst et al., 2000; Zemel et al., 2010). For cell culture, the growth area is usually homogeneously coated with ECM components. While this approach is sufficient in most cases, it does not accurately mimic most *in-vivo* niches where the ECM is highly structured (Ziółkowska et al., 2011). Microfabrication techniques enable the precise patterning of surfaces, as discussed in section 1.1.4.3.

Cells can also sense external forces such as flow shear stress through a mechanism known as mechanotransduction, a process whereby physical cues are converted into biochemical signals (Jaalouk and Lammerding, 2009). Mechanotransduction was shown to mediate cell behaviours such as proliferation, differentiation, and apoptosis. This has practical consequences for cell culture, as forces applied during routine handling of cells might alter their phenotype (Veraitch et al., 2008). The effects of physical forces on cell function and phenotype, with a focus on flow-induced forces, are extensively studied using microfluidic culture devices (see section 1.1.4.3).

1.1.4 Culture systems for anchorage-dependent cells

The primary role of the vessel is the compartmentalisation of the culture, acting as a physical barrier against potential airborne germs and other contaminants. The internal surface of the vessel often serves as growth support for the cells, and must thus be compatible with their requirements for attachment and spreading. While early culture vessels tended to be made out of glass, tissue culture polystyrene (TCPS) has become the material of choice as it enabled a shift towards pre-sterilized disposable culture vessels. The hydrophobic properties of TCPS result in the absorption of extra-cellular matrix (ECM) proteins and thus facilitate the attachment, spreading, and proliferation of most adherent cell lines (Shimizu et al., 2003). Regardless of the material, certain cell lines can only be cultured if the growth surface is coated with the appropriate ECM proteins prior to seeding.

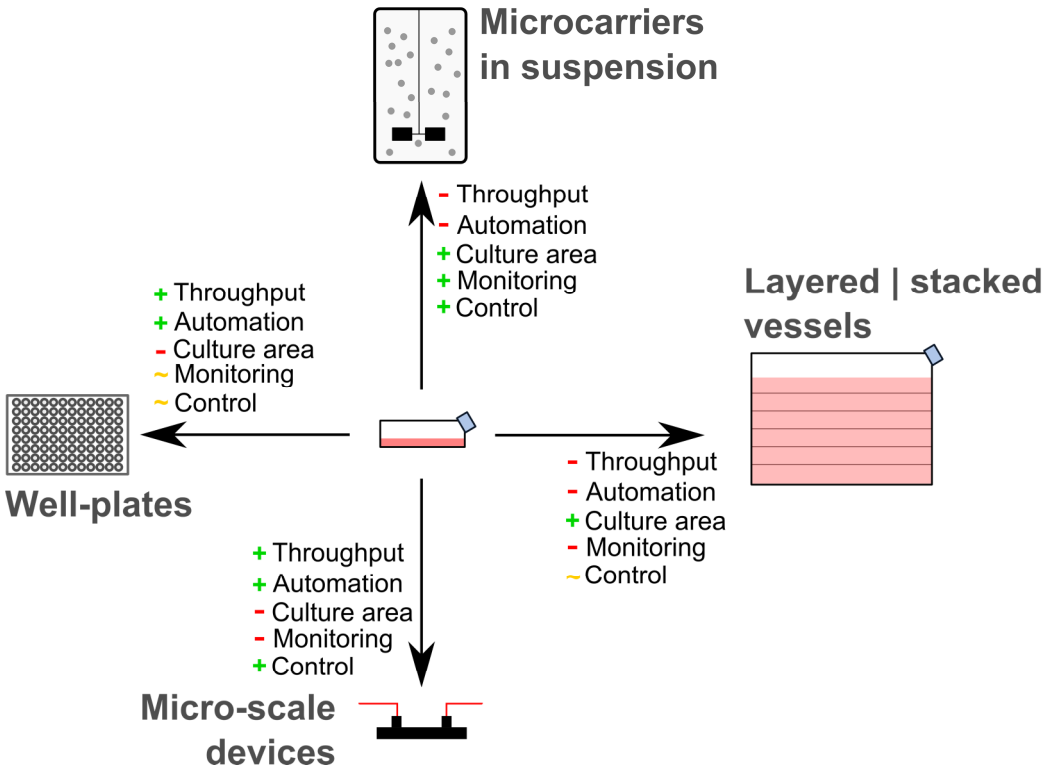


Figure 1-3 Examples of different systems for adherent cell culture. A conventional culture flask is used as baseline comparison. For each aspect compared (e.g. throughput, automation), + signifies more or improvement, - signifies less or decrease, and ~ signifies similar.

A wide range of sizes and types of vessels exists for adherent cell culture and the best choice for a specific application depends on the requirements in terms of total growth area, throughput, control, monitoring, and automation (Figure 1-3). The monitoring approach described in this work aims to be applicable to as many systems as possible, thus enabling direct comparison of experimental outcome across different scales. It was thus critical for this project to investigate different adherent cell culture approaches and identify potential aspects that can be improved with regards to their analytical and monitoring capabilities.

1.1.4.1 Flasks for routine maintenance

The vessels used for routine maintenance of cell cultures have not significantly changed since the inception of aseptic culture techniques and that of the petri dish. The T-flask is used almost ubiquitously across all fields of biology that rely on adherent cell culture (Figure 1-4). They have a rectangular footprint with effective culture areas usually varying from 25 cm^2 to 125 cm^2.

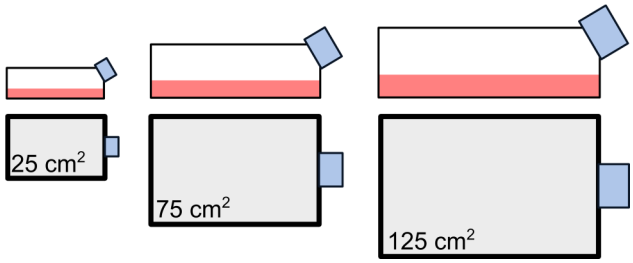


Figure 1-4 Examples of T-flasks used for routine maintenance of adherent cells.

An angled neck facilitates the access to the culture surface for medium exchange and cell handling. The cells can be imaged with any microscopy method compatible with plastic materials (i.e. most light and fluorescence microscopy methods). T-flasks represent a good middle ground in terms of culture area, and usually are sufficient to generate the number of cells required for most experiments. However, they are not suitable for experiments involving a very large number of cultures in parallel due to their large footprint and unwieldy handling. For the same reasons, T-flasks also do not lend themselves well to automation, with the exception of advanced systems such as the ones offered by TAP biosystems[1] that employ complex robotics approaches to mimic the handling of human operators. For automated cell culture, it is often preferable to use small scale culture devices such as well-plates or microfabricated culture devices instead (see sections 1.1.4.2 and 1.1.4.3, respectively). Conversely, when a very large number of cells is required, scale-up systems such as layered flasks or microcarrier cultures might be considered (see sections 1.1.4.4).

[1] http://www.tapbiosystems.com/ (last accessed 01.05.2014)

1.1.4.2 Well-plates for parallelised experiments

The design of well-plate culture vessels is derived from that of microtiter plates commonly used for small-scale chemical reactions. The biggest differences being the material used (TCPS) and flat-bottomed wells to allow cells to settle and attach (Figure 1-5). The main advantage of well-plates over T-flasks is that they can accommodate multiple culture experiments in parallel in a small footprint while using small volumes of potentially costly reagents. As they are fabricated using transparent polymers, each well can be addressed individually using light or fluorescence microscopy. As the footprint of a well-plate is standardised, this type of vessel can be used with a large number of analytical instruments that accommodate this format, such as fluorescence plate readers or image cytometers. In addition, it is readily compatible with automated liquid handling systems.

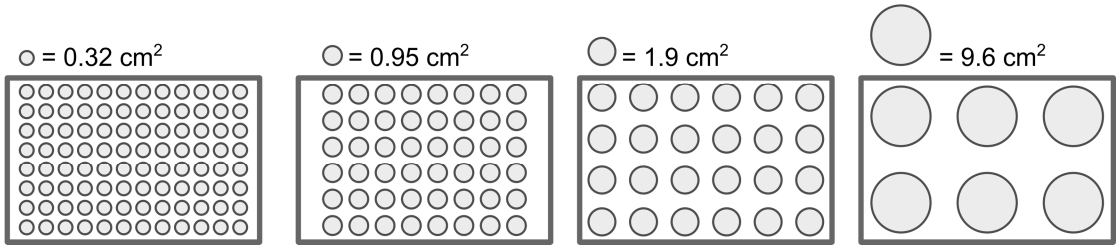


Figure 1-5 Examples of well-plates used for parallelised cultures of adherent cells. From left to right: 96, 48, 24, and 6 well-plates. The area of a single well is shown above each schematic.

Well-plates can exhibit an edge effect, whereby the wells closer to the edge of the plate are exposed to different environmental conditions that those closer to the centre, which might lead to reproducibility issues (Lundholt et al., 2003). Moreover, small wells are less suitable for microscopy inspection due to the meniscus effect, which can significantly degrade image quality.

Many experiments carried out in larger vessels (e.g. T-flasks) can be down-scaled and parallelised using well-plates without requiring significant alterations. However, the limitations remain mostly the same, with biochemical environment control limited to infrequent culture medium exchanges. Recent developments in micro fabrication techniques enable the design of devices that combine the advantage of the small form factor with capabilities such as continuous perfusion. These so-called microfluidic cell culture devices are discussed next (section 1.1.4.3).

1.1.4.3 Microfluidic cell culture devices

Microfluidic cell culture devices combine miniaturisation with the ability to precisely control fluid flow at or below the micro-scale (Mehling and Tay, 2014). This approach presents multiple advantages when compared to other scale-down culture vessels (e.g. well-plates) and T-flasks. The most obvious one is the ability to control the concentration of nutrients and metabolites by the way of continuous or intermittent medium exchanges. Rapid changes of

environmental conditions are more easily achieved in small volumes (Zhang and van Noort, 2011). Moreover, miniaturisation allow parallelizing adherent cell culture experiments, either by housing a large number of culture chambers in a single device or by scaling-out and operating multiple devices concurrently (Chin et al., 2004; Reichen et al., 2013).

Microfluidic cell culture devices enable experimental investigations that are not easily performed or not achievable with larger systems. The most evident applications are ones that can benefit from precisely controlled fluid flow. For instance, microfluidics devices can be used to study the effects of flow-induced shear forces on cell function, for example in the context of embryonic stem cell differentiation (Figallo et al., 2007; Kim et al., 2006; Toh and Voldman, 2011; Villa-Diaz et al., 2009). Likewise, it allows devising advanced medium exchange strategies (e.g. continuous or intermittent) to either favour the accumulation of secreted factors or the establishment of a quasi-steady state in terms of concentration of biochemical components (Giulitti et al., 2013; Kamei et al., 2009; Korin et al., 2009; Yoshimitsu et al., 2013). The ability to generate well-defined spatial chemical gradients within micro channels was used for the systematic and quantitative investigation of both bacterial and mammalian cell chemotaxis (Ahmed and Stocker, 2008; Kim and Wu, 2012; Nam et al., 2007). Using similar techniques, quorum sensing was investigated and manipulated in microfluidic culture devices, allowing to experimentally re-create complex predator-prey ecosystems (Balagadde et al., 2008; Janakiraman et al., 2009).

Modern microfabrication techniques enable the design of devices with features whose size is comparable to that of cellular structures (Ziółkowska et al., 2011). Artificial cellular niches can thus be engineered. For instance, micropatterning approaches can be used to generate arbitrary extracellular matrix patterns and cell arrangements. For example, gradients of ECM components were used to guide the growth and development of neurons (Millet et al., 2010). Similarly, different cell types can be co-cultured using micro patterning approaches by defining regions that selectively promote the attachment of certain cell types, thus allowing the creation of multicellular tissues akin to what is observed *in-vivo* (Khetani and Bhatia, 2008).

Another advantage that comes with miniaturisation is the ability to design devices that can accommodate a large number of cultures carried out in parallel while maintaining a small footprint. Numerous designs have been fabricated, for example 36 chambers for 3D cultures of cancer cells (Wu et al., 2011), 256 chambers for fluorescence-based gene expression analysis of rat liver cells (King et al., 2007) or 576 chambers to investigate the response of various cell lines to toxins (Wang et al., 2007). By combining such highly parallelised designs with the aforementioned capabilities (e.g. predictive flow patterns), microfluidics devices are ideal for drug discovery and toxicity studies (Neuži et al., 2012; Vladisavljević et al., 2013).

In general, microfluidics cell culture devices enabled advances in three distinct areas: control of the cells' microenvironment, high throughput experiments, and integration of further functionalities that complement the ability to grow cells (Young and Beebe, 2010). A microfluidic device accommodating all steps of a cell culture protocol (i.e. seeding, culture, and passaging) is an impressive example of the level of integration achievable (Barbulovic-Nad et al., 2010).

In many cases, microfluidic culture devices are designed so that their culture chambers can be imaged using either light or fluorescence microscopy. Indeed, such techniques are often the only practical approaches for the characterisation of cultures in microfluidic devices, as the low number of cells can preclude the use of methods that require cell detachment prior to analysis. In many cases, it is possible to compare experimental outcomes between microfluidic devices and other scales (e.g. well-plates or flasks) based on microscopy images. This aspect was a key motivation for the development of the automated microscopy image processing methods described in this work.

1.1.4.4 Roller bottles, stacked vessels and microcarrier-based culture systems

It is often necessary to scale-up the production of cells, for example to prepare a cell bank, to perform a large number of experiments, to produce recombinant therapeutics (e.g. vaccines, antibodies) or for cell therapy applications. In contrast to suspension culture, where production capacity can be expanded by increasing the volume of the reactor, geometric scale-up of vessel sizes is not suitable for adherent cultures (Want et al., 2012).

While it is possible to scale-out instead of up (i.e. use of a large number of flasks), this leads to a rapid increase in both total culture footprint and handling requirements, rendering this approach often unpractical. Roller bottles are a scale-up strategy for adherent cell processes that has been widely used for decades in the biotechnology industry with a focus on the production of vaccines (Hu et al., 2008) and more recently for tissue engineering applications (Sutherland et al., 2005). Cells are attached to the inner surface of bottles that are slowly rotated. Rolling bottles have larger effective growth areas than flasks (up to ~1700 cm^2) and the slow agitation from the 'rolling' ensures good homogeneity of the culture medium and good gas-liquid transfers (e.g. for the oxygenation of the cells). But it only slightly reduces the required footprint and amount of handling for large-scale cultures. The curvature of the bottle also makes it challenging to image the cells using light microscopy approaches.

An alternative solution is the use of layered (or stacked) culture vessels (Albeck et al., 2010). These vessels drastically increase the effective growth area (up to 5000 cm^2 or more) for a constant footprint but potentially at the expense of the homogeneity of the environment across

the different layers and that of the monitoring capabilities, as internal layers are not addressable by conventional microscopy techniques.

Microcarrier-based culture systems are a viable approach for applications where the number of cells is more critical than the ability to form well-controlled and structured cell monolayers, such as vaccine production or the expansion of pluripotent stem cells (Alfred et al., 2011; Chen et al., 2011; Hu et al., 2008). Adherent cells grow in attachment to beads (i.e. microcarriers) that are themselves kept in suspension in continuously agitated bioreactors. This allows taking advantage of instruments and methods already available for suspension cultures, such as continual monitoring and control of culture conditions as well as process scalability (Doran, 2013). It is even possible to employ *in-situ* microscopes to monitor cell growth during microcarriers cultures (Rudolph et al., 2008).

While microcarrier-based culture systems are a valuable approach for the large-scale production of anchorage-dependent cells, they represent a drastic departure from conventional adherent cell culture paradigms and were thus out of scope for this work. Instead, the focus will be put on processes that rely on standard adherent cell culture methods in microfluidic bioreactors, well-plates, flasks, and to a certain extent, stacked vessels.

1.1.5 A quantitative adherent cell culture framework

With the emergence of industrial biotechnology in the second half of the 20th century, and more recently with that of synthetic and systems biology, the ever-increasing complexity of the processes investigated required the development of novel, systematic, and quantitative methods that complement conventional experimental biology. This transition is well under-way at the level of the culture vessel itself, with new approaches such a microfluidic culture devices that enable a large palette of experiments that could not be carried out using well-established culture systems such as flasks or well-plates.

Likewise, other aspects of cell culture must be adapted to better accommodate the study of complex biological systems. For example, experimental planning based on intuition and subjective considerations can be replaced by design of experiment, a statistical framework that can help plan experiments so that their outcome will be as informative of the process studied as possible (Fisher, 1971; Fisher, 1992). Design of experiment relies on concepts such as randomisation of experiments to reduce systematic bias, replication to reduce error variance, and factorial designs where multiple parameters are varied concurrently to capture interactions (Antony, 2003).

However, all these efforts rely on the ability to generate suitable data from adherent cultures. The different methods currently available for culture monitoring will be discussed in the following section (1.2).

1.2 Non-invasive monitoring of live adherent cell cultures

The shift in biology towards quantitative methods relies on the availability of suitable analytical techniques. The ideal method would provide high quality data (i.e. quantitative information covering a wide variety of cell properties) without incurring any disruption to the culture. The latter point is key as the ability to carry out repeated measurements of a single culture allows generating the time course data that is essential to understand biological systems that are, by nature, very dynamic. However, in practice there exists a trade-off between the quality of the data generated and the invasiveness of the method (Figure 1-6).

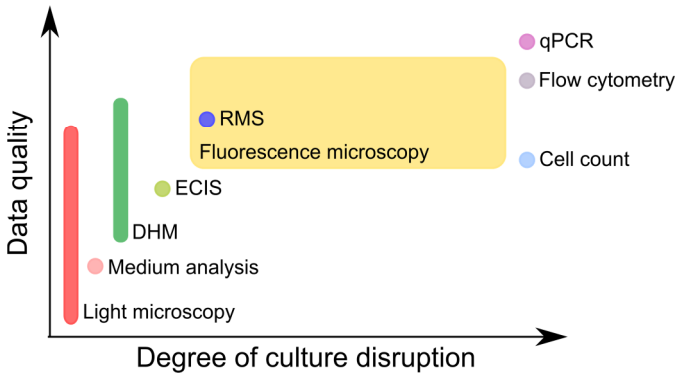


Figure 1-6 Data quality and disruption for a few adherent cell culture analytical techniques. Data quality encapsulates the robustness of the measurements (i.e. qualitative or quantitative) as well as the variety of cell properties that can be measured using the method (e.g. population size, gene expression, morphology). "DHM" stands for digital holographic microscopy, "RMS" for Raman microspectroscopy, "qPCR" for quantitative polymerase chain reaction, and "ECIS" for electric cell-substrate impedance sensing.

Usually, the gold standard methods for adherent cell culture characterisation (e.g. qPCR or flow cytometry) are highly invasive in that they require cells to be detached prior to measurements. In contrast, non-invasive methods tend to generate qualitative data (e.g. visual inspection of culture vessels through light microscopy), unless equipped with automated methods such as image analysis. As such, some methods cover a large portion of the 'quality data' spectrum depending on the variant used. Similarly, the degree of disturbance will also depend on the specific protocols employed. The type and quality of data desired will thus dictate which method to choose for a specific application (Table 1-1).

This section will first give a brief overview of invasive and destructive methods that are usually considered the gold standard for cell culture characterisation, before detailing non-invasive methods such as light microscopy, digital holographic microscopy, fluorescence microscopy, Raman microspectroscopy, impedance sensing, and indirect monitoring.

Table 1-1 Non-exhaustive summary of analytical methods for adherent cell culture characterisation

Method	Standard culture vessels	Detachment needed	Qualitative	Quantitative	Continual monitoring	Individual cellular objects	Population
Light microscopy (operator)	×		×			×	×
Light microscopy (automated)	×		×	×[1]	×[2]	×	×
Fluorescence microscopy			×	×[1]	×[2]	×	×
Digital holographic microscopy	×		×	×	×[2]	×	×
Raman microspectroscopy				×	×[3]		
Electric cell-substrate impedance sensing				×	×		×
Cell counting	×	×		×			×
Quantitative polymerase chain reaction	×	×		×			×
Flow cytometry	×	×		×		×	×

1. Using automated image processing software
2. Requires on-stage incubator or incubator with integrated microscopy capabilities
3. Only short-term monitoring has so far been reported

1.2.1 Methods requiring detachment of the cells

A large palette of analytical methods for cell cultures was developed with suspension systems in mind. Indeed, most of the gold standard approaches requires a solution of free-floating cells. Thus when working with adherent cultures, cells have to be detached prior to analysis. These destructive methods (i.e. methods that negatively impact the viability, function or structural integrity of the culture) are not suitable for the analysis of characteristics that are only observable when cells are attached (e.g. cell morphology). Measurements are often carried out when cells are passaged using a surplus of cellular material. In order to determine cellular kinetics, multiple cultures are seeded simultaneously and are sacrificed for analysis at various time points. This method relies on the assumption that the cultures can be considered identical, which is often not valid in practice due clonal heterogeneity and the stochastic nature of cell processes as well as small variations in experimental conditions (e.g. edge effect in well-plates). Averaging measurements across non-identical populations might thus prevent the detection of changes in cell properties or behaviours that are small in comparison to inter-sample variations. Nonetheless, destructive methods remain ubiquitously used and set the standard for the development of novel non-invasive monitoring approaches.

1.2.1.1 Cell counting

Cell counting is among the first procedure taught to new cell culture researchers. This is traditionally done using an haemocytometer, a device originally designed to enumerate red

blood cells (Strober, 2001). The counting chamber has a known volume and is patterned with squares of a defined area. A cell suspension is pipetted in the chamber, which is sealed using a microscopy coverslip. Using a light microscope, the cells in each square are enumerated. Based on a simple formula, the cell density of the original solution can be calculated. The same approach can be used to estimate the viability of cells by mixing the solution with the Trypan blue dye before counting. Trypan blue permeates into non-viable cells (whose membrane integrity has been compromised) and stains them blue. Viable and non-viable cells can then be enumerated separately, enabling the calculation of the fraction of viable cells. This approach is quick and cost-efficient but typically has significant inter-operator variability. As an alternative, numerous instruments have been developed to automate the process, thus increasing the robustness of the measurements (Dodet et al., 2010).

1.2.1.2 Flow cytometry

Flow cytometers allows the analysis of individual cells of a cell suspension (Bendall et al., 2012; Fulwyler, 1965). Using a principle termed hydrodynamic flow focusing, the sample solution containing the cells is forced into a narrow stream so that cells are well separated and can be interrogated individually (Figure 1-7). A light source (usually a laser) is aimed at the sample stream, often at a direction that forms a 90° angle with the direction of the stream. As cells flows past the beam, they will interfere with it. A first detector is placed directly in line with the laser beam to measure the forward scatter (FS) induced by the interrogated cell. FS is a measure of the diffracted light and informs on the cell volume. Other detectors are placed at an angle of approximately 90° to measure the side scatter (SS). SS is a measure of the refracted light, and thus informs on the internal composition of the cells. Based solely on those signals, it is possible to quantify cell size and internal complexity, enabling the determination of distinct subpopulations within a sample.

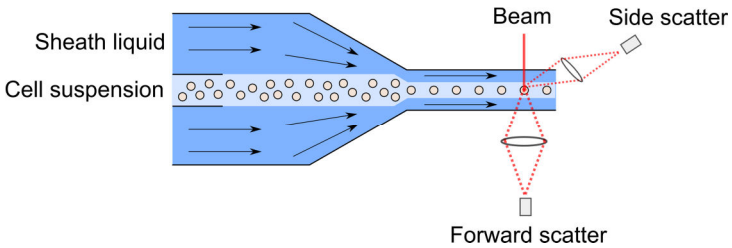


Figure 1-7 Schematic of a flow cytometer. Hydrodynamic flow focusing by a sheath liquid is used to create a stream containing well-separated cells. A laser beam is used to interrogate the cells. Forward and side scatter are measured by detectors placed in-line with the beam and perpendicularly to the beam, respectively.

In addition to FS and SS, flow cytometers usually enable the detection of fluorescence signals. Multiple lasers, filters and detectors enable the excitation of fluorophores at various wavelengths and the measurement of the emitted light. The same labelling methods than those for fluorescence microscopy apply here (see section 1.2.4). For example, cell can be labelled using fluorescent antibodies specific for surface markers (Darzynkiewicz et al., 2011).

Intracellular fluorescence proteins such as GFP can also be detected using flow cytometry (Veraitch et al., 2008). Similarly, small molecule dyes can be used to determine the viability of a population or the proportion of cells at a given stage of the cell cycle (Pozarowski and Darzynkiewicz, 2004).

1.2.1.3 Molecular biology approaches

After detachment, cells can be lysed and their content recovered for further analysis. There is a wide range of molecular biology methods that can be employed to carry out such analyses. Western blots is a techniques that enables the detection of protein from cell extracts (Burnette, 1981). It combines an electrophoretic step where proteins are separated based on their properties (e.g. isoelectric point, molecular weight) with a detection step based on immunochemistry (specific reaction of antibodies with their target antigens) as described later for fluorescence microscopy (see section 1.2.4). Similarly to its use in fluorescence microscopy, the indirect immunochemistry approach results in an amplification of the signal, which can enable the detection of protein present only in small quantities.

Polymerase chain reaction (PCR) revolutionised biology by enabling the rapid *in-vitro* amplification of DNA fragments, which truly kick started the modern-era of molecular biology (Saiki et al., 1988). A related technique termed reverse transcriptase PCR (RT-PCR) can be used to amplify messenger RNA (mRNA) and thus informs on the genes that are actively transcribed by the cells (Bustin, 2000). An additional preliminary step consists in using a reverse transcriptase enzyme to synthesise complementary DNA from mRNA, which can then be analysed by regular PCR. Quantitative results can be obtained by quantitative PCR (qPCR) where the amount of product obtained at each cycle of amplification is determined using fluorescent dyes (Ponchel et al., 2003).

1.2.2 Bright field and phase contrast methods

Light microscopy played a key role in the development of modern cell culture protocols. Indeed, it enabled the observation of cells without any disruption to the culture (Murphy et al., 2013). Conventional microscopy employs bright field illumination to detect objects that are capable of altering the transmission of light (e.g. through scattering or absorption). Those objects, such as cells or stained tissue sections, are referred to as amplitude specimen (Figure 1-8.B). While unstained live adherent cells are often transparent and thus only slightly alter light amplitude, they produce a significant shift in the phase of the light wave exiting the sample. Indeed, if the refractive index of the surrounding media and that of the specimen differ, the length of the optical path will increase or decrease proportionally to the thickness of the specimen, resulting in a phase shift (Figure 1-8.C). However, phase shifts are not perceivable by either human eyes or imaging sensors.

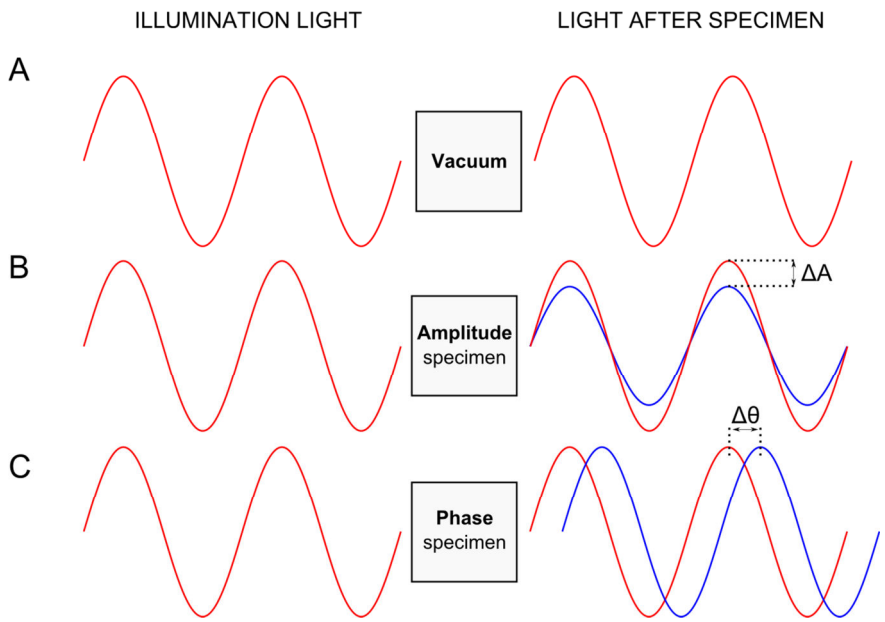


Figure 1-8 Light interaction with specimen. Interaction of the light wave front with (A) vacuum, (B) an amplitude specimen, and (C) a phase specimen. ΔA and Δθ indicate changes in light amplitude and phase, respectively.

Phase contrast microscopy (PCM), a technique pioneered by Frits Zernike (Zernike, 1942), solved that issue by translating phase variations into variations of amplitude. Its working principle is based on the separation of the background wave front from the one that was diffracted after travelling through the specimen (Figure 1-9.A). The latter is usually retarded by 90°, or approximately a quarter relative to the un-diffracted wave front (Otaki, 2000). A phase ring positioned at the objective rear focal plane is used to retard the background, un-diffracted light by 90°. Because it is positioned at the conjugate focal plane of the annular aperture, most of the background light will be phase retarded. The phase ring is made up of both phase retardant material as well as a neutral density filter, effectively reducing the amplitude of the background wave front (Figure 1-9.B). Past the phase ring, both the diffracted and background wave fronts are in-phase and will interact at the image plane. The intensity of the resulting constructive interference pattern will be dictated by the amount of light that did not travel through the phase ring and will therefore be proportional to the phase shift induced by the specimen. It effectively translated changes in phase to observable changes in intensity. Whereas adherent cells are hardly visible in bright field microscopy, cellular structures such as the membrane and intracellular compartments are clearly discernible when using PCM (Figure 1-10).

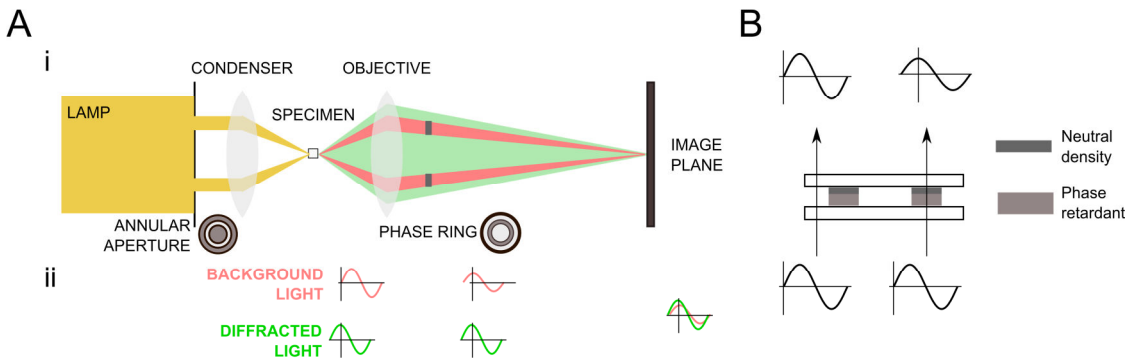


Figure 1-9 Working principle of phase contrast microscopy (PCM). (A) Schematic of the optical configuration of a phase contrast microscope (i) and the wave front patterns at various location of said configuration (ii). (B) The phase ring and its effect on incident light. Both figures were inspired and partly adapted from (Holzner et al., 2010; Ruzin, 1999; Zernike, 1942)

PCM also has the added advantage of relying on unpolarised light illumination. Indeed, a number of other contrast enhancing techniques, such as differential interference contrast (DIC) microscopy, employ polarised light to enable the observation of phase specimen. In principle, DIC can produce images of very high quality that are mostly artefact-free. However, it is not suitable for the imaging of cells cultured on optically active (or birefringent) materials, which alter the propagation of polarised light and result in significantly degraded image quality (Cui et al., 2008). Because tissue culture polystyrene is a birefringent material, PCM is often preferred for adherent cell culture imaging.

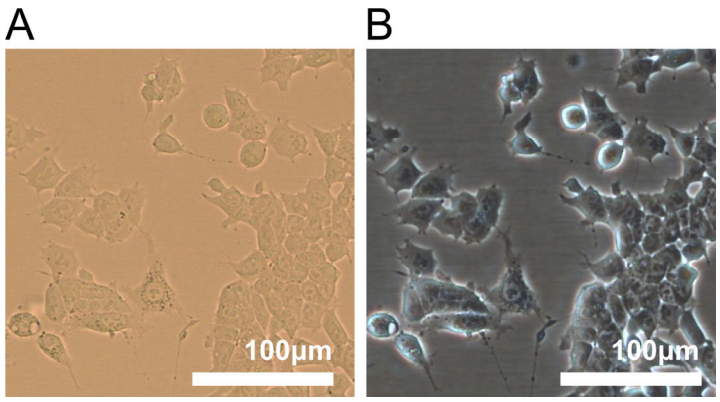


Figure 1-10 Comparison of bright field and phase contrast microscopy. Bright field (A) and phase contrast (B) microscopy images of the same field of view of a mouse embryonic stem cell culture.

There are certain drawbacks to PCM that are intrinsic to the method and the underlying image formation process. A small fraction of the light diffracted by the specimen can still pass through the phase ring, together with the un-diffracted background light. This small fraction of the specimen light, mostly corresponding to low spatial frequencies (i.e. light diffracted at small angles), will thus remain phase shifted relative to the background light by 90° and be attenuated by its passage through the neutral density filter (Murphy et al., 2013). The interference pattern at the image plane is not destructive and therefore results in a local inversion of the contrast, taking the form of bright halos artefacts. These are most noticeable for large, low spatial frequency objects such as intracellular compartments (e.g. nuclei) or even whole cells in some cases (Murphy et al., 2013). These halos can potentially obfuscate the cell membrane and thus

make it difficult to determine the true shape of cellular objects. The severity of the halo artefacts can be partially decreased by using a medium with a higher refractive index, though this solution is often not practical when observing live cells. Alternatively, halo effects can be attenuated using an apodized objective where two concentric neutral density filters surround the phase retardant material. Their role is to reduce the intensity of light diffracted at small angles by low spatial frequency objects (Otaki, 2000).

Another common PCM artefact that is related to halos is the shade-off (or zone-of-action) effect. It occurs when the intensity of the image is not directly proportional to the phase shift produced by the specimen. Whereas one could expect a uniform specimen to produce a constant intensity along its length, it is rarely the case when observing adherent cells using PCM. Indeed, at the centre of a large specimen, the angle of diffraction is much less than that at its edges, which are highly refractive regions (Murphy et al., 2013). Light from the centre of the specimen is likely to be captured together with background light by the phase ring. As a result, the centre of the object and the image background will have very similar intensities. The intensity profile in the neighbourhood of a cell on a PCM image is thus shaped by both the halo and shade-off effects (Figure 1-11).

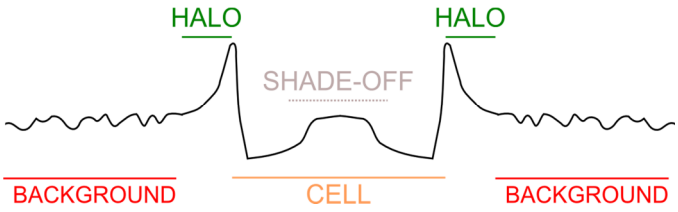


Figure 1-11 Typical intensity profile in the neighbourhood of a cell on a PCM image. The background intensity shows slight fluctuation around a stable mean. As it approaches the cell, the intensity soars due to the bright halo artefact before quickly decreasing as it enters the darker region corresponding to the cell. Due to the shade-off effect, the intensity steadily increases until it reaches levels close to that of the image background.

Light microscopes were conventionally built in an up-right configuration, with the light source and the objectives above and below the specimen, respectively. While this configuration works well with sample preparations such as microscope slides, it doesn't accommodate cell culture devices due to their thickness, resulting in a distance between the objective and the cells attached to the bottom of the vessel that is greater than the working distance of most objectives. As an alternative, inverted microscopes have the specimen illuminated from the top with the objective placed beneath it (Lawrence, 1852). This way, the distance between the objective and the cells is well within the specification of most objectives and thus cells in culture vessels can be observed directly.

In many cell culture laboratories, the microscopy equipment is limited to a simple inverted phase contrast microscope without any automation or image acquisition capabilities (Figure 1-12.A). This setup is suitable for visual inspection of cultures but estimation of cell culture characteristics would have a very high degree of uncertainty due to operator bias and

subjectivity. Such a setup should not be used for decision-making or evaluation of experimental outcomes. The addition of a digital camera for image acquisition enables proper documentation of the observations and the use of image analysis methods for the generation of quantitative data (Figure 1-12.B). As the imaging environment is not controlled, image acquisition should be as quick as possible in order to minimise the amount of time during which cells are left in a sub-optimal environment. As the culture vessels have to be transported between the incubator and the microscope for each analysis, the time between data points is usually significant, especially if multiple cultures have to be monitored concurrently. Another possible microscope configuration includes an automated stage as well as an on-stage incubator that enables imaging of cells in an optimal culture environment (Figure 1-12.C). Image acquisition can be fully automated and thus much more frequent. It also facilitates time-lapse microscopy whereby the same fields of view are imaged over time, allowing the tracking of individual cellular objects and the characterisation of their behaviour (e.g. migration, morphological changes or proliferation). It is the ideal setup for adherent cell culture monitoring but it also limits the number of cultures that can be monitored as such configuration can only accommodate one or at best a few cell culture vessels simultaneously.

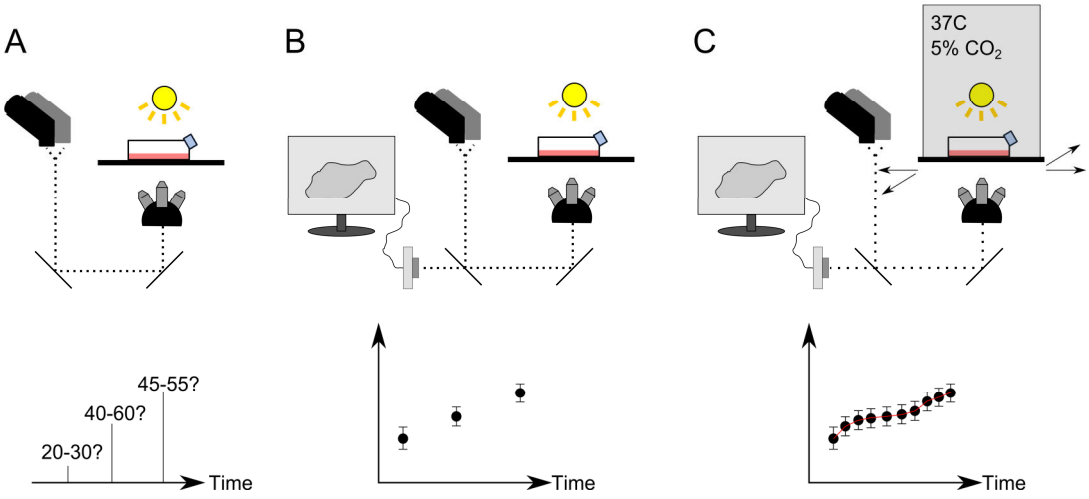


Figure 1-12 Three different phase contrast microscopy setups. (A) Manual setup, (B) a setup with a digital camera for image acquisition and (C) an automated setup with a motorised stage, a camera for image acquisition, and an on-stage incubator for environmental control. The second row illustrates the type of data that can be generated form these systems: (A) only qualitative with high uncertainty, (B) quantitative (image analysis) with low uncertainty but intermittent sampling, (C) quantitative (image analysis) with low uncertainty and frequent sampling.

Specialised systems have been developed to alleviate the throughput issues of live cell microscopy. Incubators with integrated microscopy capabilities include the Nikon BioStation CT (Herbert et al., 2009) and CM Technologies' Cell IQ (Narkilahti et al., 2007). Alternatively, another approach consists in placing a stand-alone imaging unit in a conventional incubator. Essen BioScience developed Inucyte, an imaging system compatible with the atmosphere of an incubator (Pierscionek et al., 2012). The LumaScope from Etaluma has similar capabilities (Gouveia et al., 2013). These solutions are attractive but remain expensive. In addition, these

systems do not yet offer the versatility of a traditional light microscopy setup, which remains the gold standard for non-invasive monitoring of cell cultures.

1.2.3 Digital holographic microscopy

Digital holographic microscopy (DHM) is a recent advance in light microscopy, and is part of a larger group of methods termed quantitative phase contrast microscopy (Marquet et al., 2013). The light of a coherent light source (usually a laser) is split into an object beam and a reference beam. The object beam travels through the specimen and is collected by a conventional microscope objective before interfering with the reference beam to form a hologram, which in this case is defined as the interference pattern between two mutually coherent waves (Kühn et al., 2013). In essence, the resulting hologram encodes the amplitude and phase modulations induced by the specimen. Using this encoded data, numerical methods based on the theory of diffraction can be used to simulate the propagation of the wave front, enabling the reconstruction of an in-focus intensity and phase image (Marquet et al., 2013). The software essentially replaces the lens of a conventional microscope. A simpler configuration was proposed in the form of digital in-line holographic microscopy (DIHM) where the reference and object spherical waves are created by placing a pinhole in front of the light source. Both waves will interact to form the hologram (interference pattern) at the plane of the detector of a camera situated directly behind the specimen (Xu et al., 2001).

The hologram captured by DHM encodes the 3D information of the specimen, which means that focusing can be done after the fact during in-software reconstruction (Mölder et al., 2008). Moreover, because image formation is done through numerical reconstruction instead of relying on expensive and fairly large optics, compact and affordable DHM instruments could be developed (Seo et al., 2009). New iteration of commercial instruments can even fit into a conventional cell culture incubator. Image acquired using DHM can, at least in theory, have a quality close to that of fluorescence microscopy images (i.e. high contrast between cells and image background), which greatly simplifies their analysis using image processing methods (Kühn et al., 2013). However, the potentially cytotoxic effect of the laser illumination (which can be in principle be alleviated by the use of LEDs as light sources) and the requirement for entirely new instruments might slow down the adoption of this promising technology.

1.2.4 Fluorescence microscopy

Fluorescence microscopes are conceptually very similar to their light counterparts but instead of relying on the interaction of a light wave front with the specimen, it is based on the phenomenon of fluorescence. Electrons of a fluorophore molecule absorb high energy photons (if resonant, i.e. of energy equal to the transition energy between ground state and excited state)

and are excited to a higher quantum state before relaxing back to their ground state, a process during which a photon is emitted (Lakowicz, 2010). The emitted photon will usually be of lower energy (thus of a higher wavelength) than the excitation photon, a phenomenon known as the Stokes shift.

The advantages of fluorescence microscopy over light microscopy include a better contrast between the background and the fluorescent specimen as well as the possibility to selectively label structures of interest (Fritzsche and Mandenius, 2010). In most cases, fluorescence microscopy is compatible with existing light microscopes but requires additional components: a suitable light source, an excitation filter, an emission filter and a digital camera (Figure 1-13.A). Such a setup enables wide-field fluorescence imaging, where the entire sample volume is evenly illuminated by the excitation light. This is in opposition to laser-scanning confocal microscopy that enables the targeted point-illumination of structures in the focal plane. Confocal microscopy is discussed in more details below.

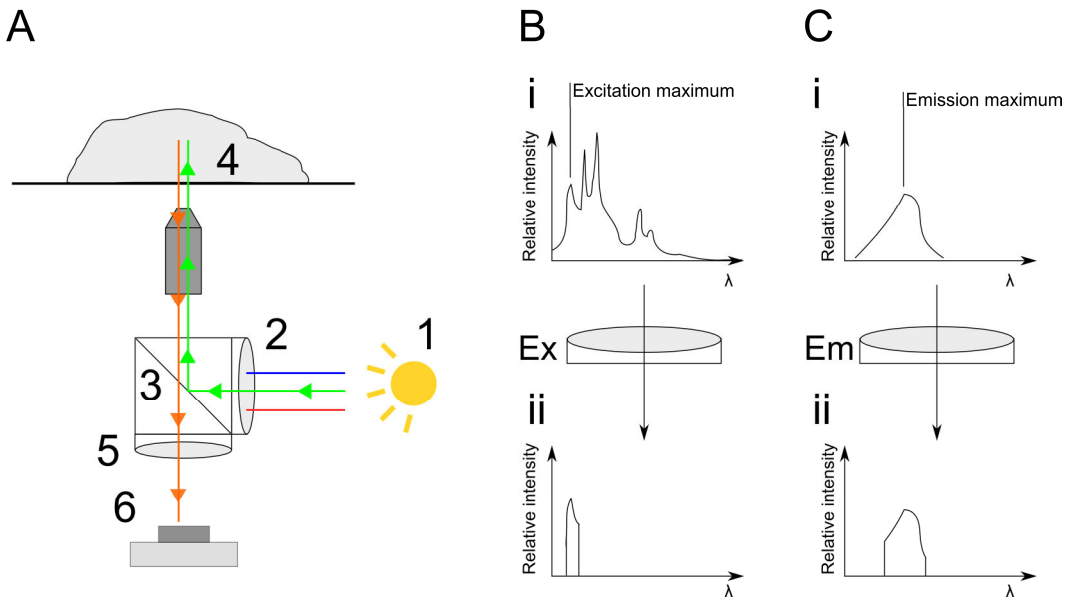


Figure 1-13 Wide-field fluorescence microscopy principles. (A) Diagram of a typical wide-field fluorescence microscopy setup. (1) Light source, (2) excitation filter, (3) dichroic mirror, (4) specimen, (5) emission filter, (6) digital camera. (B) Relative excitation intensity in function of the wavelength before (i) and after (ii) the excitation filter for a typical mercury lamp. (C) Relative emission intensity in function of the wavelength before (i) and after (ii) the emission filter.

The conventional approach to illumination for fluorescence microscopy consists in using an excitation source that produces light of high intensity across most of the wavelength spectrum (termed multispectral light), for example a mercury or xenon lamp (Figure 1-13.B.i). The excitation light then travels through a filter so that only wavelengths within a defined band are selected while the rest is blocked (Figure 1-13.B.ii). The band of wavelengths to select is mainly dictated by the excitation spectrum of the fluorophore to be imaged. This step is essential for samples that might contain more than one type of fluorophore as those can then be selectively excited if there is no significant overlap in their excitation spectra. The filtered light is then reflected towards the objective using a dichroic mirror so that the samples can be illuminated.

The fluorophore contained within the sample will be excited and emit light at a longer wavelength (Figure 1-13.C.i). The emitted light will travel in all directions, including towards the objective. The fraction of the emitted light going through the objective can travel downwards unaffected by the dichroic mirror, as it is not sensitive to its wavelength. It passes through an emission filter, whose role is to filter out residual illumination light and undesired excitation light, for example from other fluorophores (Figure 1-13.C.ii). The light finally reaches the digital camera to form the fluorescence image, on which the contrast between the fluorophore and the background should be high given the right selection of excitation filter, dichroic mirror and emission filter. A digital camera is often required to achieve the long exposure time necessary for the observation of the fluorophore due to low emitted light intensity.

This setup can be further improved by replacing the discharge lamp by an array of light-emitting diodes (LEDs) as the excitation source. Mercury or xenon lamps can have highly fluctuating illumination intensities (especially during the warm-up period) and short lifetimes (usually less than 1000 hours for a typical bulb). In contrast, LED systems have a very stable illumination output, can be rapidly switched, and their lifetime is estimated to be ~50,000 hours (Albeanu et al., 2008; Sato and Murthy, 2012). Moreover, LEDs usually have a narrow wavelength excitation band, minimising the risk of excitation light bleeding through the filters.

As an alternative to wide-field fluorescence methods, confocal microscopy approaches have increased spatial resolution (~1.4-fold in theory) and enable optical sectioning of samples (Wilson, 2011). A single point is illuminated (usually using a laser as excitation source) and the light emitted is collected through a pinhole in an optically conjugated plane. Most of the out of focus light (i.e. light that is emitted from a region away from the illuminated point) will thus not be detected. As only a single point is imaged at a time, the specimen is scanned and an image is virtually constructed using specialised software. Confocal microscopy is not generally used for live cell monitoring due to the phototoxicity of the excitation laser, the incompatibility with common plastic-based culture vessels, and the very high cost of the necessary instrumentation (Paddock, 2000).

Another method, termed two-photon microscopy, produces similar results to that of confocal microscopy but significantly reduces the issue of phototoxicity. It is based on the phenomenon of two-photon absorption: an electron that would normally be excited by a photon at a given wavelength can also absorb two photons of approximately twice the wavelength (Denk et al., 1990). For example, a fluorophore with a single photon excitation peak at 340 nm can be excited with two photons at 680 nm. Because a high photon density is necessary to achieve two-photon excitation, only the volume directly illuminated by the laser will eventually be affected by photodamage (i.e. no out-of-focus phototoxicity), unlike confocal microscopy where the

excitation light can potentially interact and damage the material in its path (Straub et al., 2000). Two-photon microscopy is therefore more suitable than confocal microscopy for the monitoring of live cell cultures.

Regardless of the fluorescence microscopy method chosen, the cells or structures thereof have to be fluorescently labelled. There are three main approaches used for the labelling of biological samples: immunofluorescence, use of small organic fluorophore molecules, and genetically encoded fluorescence (Figure 1-14).

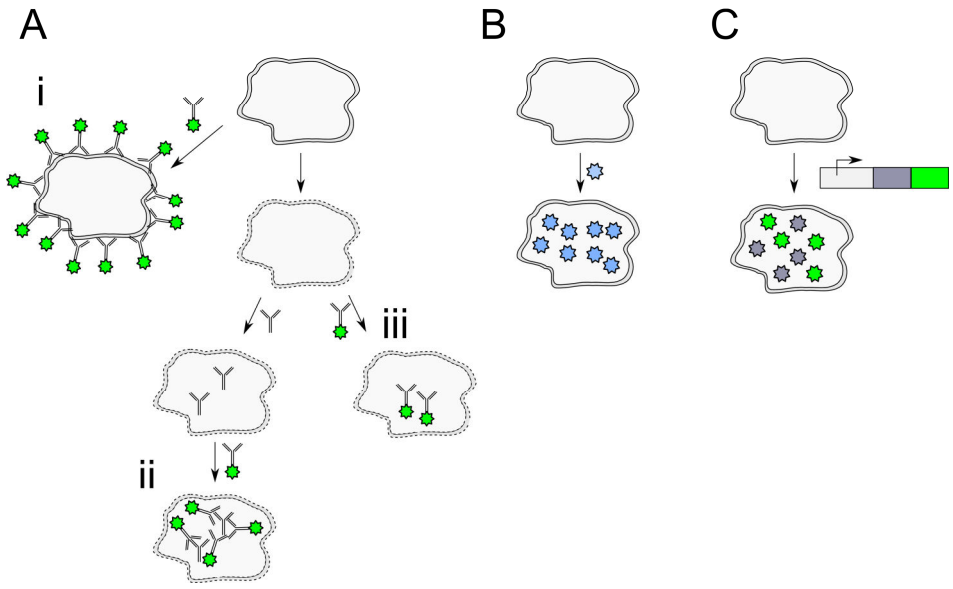


Figure 1-14 Strategies for fluorescence microscopy. (A) Immunofluorescence where antibodies are used to selectively label structures of interests. If the target is a membrane protein, cells can be directly labelled using antibodies conjugated with a fluorophore (i). If the target is intracellular, the cells must first be fixed and permeabilised. At this stage, indirect labelling can be employed by first incubating the cells with a primary antibody specific for the target structure followed by the addition of a secondary antibody (specific for the primary antibody) conjugated with a fluorophore (ii) or directly use an conjugated antibody specific for the structure of interest (iii). (B) Use of small fluorescent molecules that can diffuse through the membrane and thus be used on live cells. (C) Genetically encoded fluorescence. The gene coding for a fluorescent protein is inserted into the expression system of the gene of interest. As a result the fluorescence protein will be co-expressed with the gene of interest.

Immunofluorescence relies on the specificity of antibody-antigen binding to target cellular structures of interest (Odell and Cook, 2013; Shakes et al., 2012). Fluorophores such as fluorescein isothiocyanate (FITC) or the more recently synthesised Alexa dyes (Panchuk-Voloshina et al., 1999) can be functionalised by attachment to a succinimidyl-ester group, which in turn can react specifically to primary amines of antibodies. This process results in conjugated-antibodies that can be used to fluorescently label the corresponding antigens (Figure 1-14.A). This approach can be used with live cells if the target antigen is localised on the surface of the cells (Nethercott et al., 2011), although the extent to which this type of labelling might impact the normal function of the cells is not well characterised (Figure 1-14.A.i). In most cases though, immunofluorescence is carried out on samples that have been prepared by fixation and permeabilisation so that intracellular antigens can be labelled (Jamur and Oliver, 2010). Fixation is usually carried out using a protein cross-linking agent such as paraformaldehyde (PFA), which essentially fixes the cells in their last living state and allows for

long-term storage. Organic solvents (e.g. methanol, acetone) or detergents (e.g. Triton, Tween) can be used to permeabilise the cells so that antibodies can reach intracellular structures through diffusion.

Immunofluorescence of fixed and permeabilised samples can be direct or indirect. The direct method is similar to the previously mentioned labelling of the cell membrane. The prepared sample (fixed and permeabilised) is incubated with a fluorescently-conjugated antibody, which will diffuse into the cells and label the antigen of interest (Figure 1-14.A.ii). This one-step labelling approach is simple but can result in faint signals. Instead, it is sometime preferable to use an indirect labelling method whereby a primary (unconjugated) antibody is first used to label the target antigen before incubating the cells with a fluorescently-conjugated secondary antibody that binds specifically to the primary antibody (Figure 1-14.A.iii). Because multiple secondary antibodies can bind to a single primary antibody, the fluorescence signal is amplified, resulting in high contrast images.

As an alternative to immunochemistry, small fluorescent molecule can bind specifically to various cellular structures (Figure 1-14.B). The main advantage of such a method is that the dyes will usually diffuse freely through the cell membrane and can, in principle, be used for the labelling of live cells. DAPI (4',6-diamidino-2-phenylindole) is a fluorescent dye that binds to DNA (Zink et al., 2003). In order to stain live viable cells, high concentrations of DAPI are required, potentially resulting in cytotoxic effects. Hoechst is a dye similar to DAPI in that it will also diffuse through the cell membrane and bind to the DNA (Portugal and Waring, 1988). It is generally regarded as less toxic than DAPI and is therefore the dye of choice for nuclear labelling of live cells even though it was shown to significantly impact cell mechanisms such as cell cycle (Durand and Olive, 1982). Acetomethoxy derivate of calcein (calcein AM) is a dye that diffuses through the cell membrane and into the cytoplasm where it is cleaved by intracellular enzymes, which activates its fluorescent properties and modifies its structures so that it is less likely to diffuse out of the cell (Weston and Parish, 1990). Chloromethylfluorescein diacetate (CMFDA), chloromethyltetramethyl rhodamine (CMTMR) and carboxyfluorescein diacetate succinimidyl ester (CFDA-SE) are other dyes that behave similarly to calcein AM but with different absorption/emission spectra and intracellular lifetimes (Lantz, 2001; Lyons, 1999). While it is generally accepted that low concentration of those dyes do not significantly alter cell proliferation or viability, it was shown that they can have an effect on cell mechanics (Lulevich et al., 2009). Similar dyes are available for the labelling of various live cell components, including the cytoskeleton (Chazotte, 2010), mitochondria (Johnson et al., 1980), and cell membrane (Chazotte, 2011).

In addition to immunofluorescence and small molecule dyes, genetically encoded fluorescence has been widely used for fluorescence microscopy applications. This was first made possible by

the discovery of the green fluorescent protein (GFP), isolated from the jellyfish *Aequora victoria* (SHIMOMURA et al., 1962) followed by the determination of its coding sequence (Prasher et al., 1992) and its expression in *E. coli* and *C. elegans* (Chalfie et al., 1994). The main application of this technology, which led to a revolution in biology, is the use of the GFP as a marker for gene expression by having its coding sequence cloned in a region of the genome under the control of the same regulatory elements than the gene of interest. In other words, gene expression could be visualised, allowing the study of spatial and temporal expression patterns. There is now a large palette of fluorescent proteins that can be chosen from depending on the desired absorbance and emission spectra (Shaner et al., 2005). In addition, variants with lower half-life (from about 24 hours for the wild type to ~2 hours) were engineered to allow for the detection of short-term dynamic changes in gene expression (Li et al., 1998). While genetically encoded fluorescence is widely used, it is not without its limitations. The transduction of the GFP gene into cells is obviously an intrusive process and results in a new clone not necessarily representative of the wild type cells being studied. Both the transduction process and the overexpression of intracellular GFP were found to be potentially detrimental to cells (Goto et al., 2003; Liu et al., 1999).

1.2.5 Raman microspectroscopy

Raman microspectroscopy (RMS) is a hybrid method that combines a traditional microscopy setup with Raman spectroscopy (Notingher and Hench, 2006). Raman scattering can be related to fluorescence in the sense that both processes involve the interaction of a photon with a molecule, resulting in a transfer of energy between the two. However, while fluorescence is based on the absorption of the photon, the Raman effect is related to the scattering of the incident photons. Most scattering events are elastic and the energy of the scattered photons remains unchanged, a phenomenon known as Rayleigh scattering. In a small number of cases (~1 in 10^8), the scattering will be inelastic and a small portion of vibrational energy will be transferred to or from the photon, resulting in a Raman shift. Due to the Stokes shift previously discussed for fluorescence (section 1.2.4), the wavelength of the scattered photon will be shifted accordingly. The magnitude of the shift will be dependent on the vibrational energy of the molecule. A Raman spectrum can be constructed by determining the number of photons scattered (i.e. the intensity of the signal) for various Raman shift values (i.e. different wavelength). This spectrum constitutes a fingerprint of the sample's molecular structure (e.g. atoms, type of bonds). Whereas fluorescence has a defined lifetime, Raman scattering is nearly instantaneous.

A microscope fitted for RMS allows carrying out both traditional light microscopy and Raman spectroscopy (Notingher and Hench, 2006). A laser (usually visible or near-infrared) is used as the excitation source. A standard microscope objective is used to focus the laser beam into the

sample and to collect the scattered photons. A notch filter separates the Rayleigh photons (same energy as the photons of the excitation laser) from the Raman shifted photons in order to minimise background noise. A diffraction grating disperses the photons into multiple beams that each corresponds to a different Raman shift (i.e. wavelength) before being collected on an array of detectors. A large laser spot size can be used to illuminate a whole cell at a time (~500 μm^2) and generate a single Raman spectrum per cell (Pyrgiotakis et al., 2009). Alternatively, the size of the laser spot can be drastically reduced. The spatial resolution is limited by diffraction and is approximately half the wavelength of the excitation source, allowing the generation of spectra at a sub-micrometre resolution (Puppels et al., 1990). As such, it is possible to generate an average cell spectrum and to reconstitute images (termed Raman spectral images) where the intensity of the pixel corresponds to the high of a certain peak in the Raman spectrum (Pascut et al., 2011). However, this last approach significantly increases the amount of time required to analyse a cell as hundreds of spectra are required and each spectrum can take from a second to minutes to acquire depending on the experimental setup.

A major downside of RMS is that the analysis of Raman spectra can be challenging. Methods such as principle component analysis (PCA) or support vector machines (SVM) were used to classify the spectra (Pascut et al., 2011; Pyrgiotakis et al., 2009). In addition to the complexity of data analysis, the long acquisition time and the potential cytotoxicity of prolonged exposure to laser excitation, RMS require the use of non-standard culture substrates such as fused-silica (Notingher et al., 2004) or magnesium fluoride (Pascut et al., 2011) in order to minimise undesired background signal. Thus cells cannot be imaged in conventional culture vessels.

1.2.6 Impedance sensing

Before the advent of modern imaging methods and the development of suitable image analysis algorithms, the only non-destructive approach for adherent culture characterisation was visual inspection using light microscopy, which is intrinsically qualitative. Alternative, non-imaging based approaches were proposed for the generation of quantitative cell culture data. One such approach is electric cell-substrate impedance sensing (ECIS), a method based on the ability of cells to behave as electrical insulators (Giaever and Keese, 1993). Electrodes are patterned at the bottom of the cell culture device. The culture medium acts as the electrolyte and the cells attach and spread on the electrodes, which will impede the flow of an applied alternative current. Simple impedance, which can be related to Ohm's law for direct current, is given by the ratio of the voltage measured across the electrode to the alternative current. Complex impedance can also be measured, which has a real (resistance) and an imaginary (capacitive reactance) part (Bagnaninchi and Drummond, 2011). The capacitive reactance can be used to calculate the capacitance, which unlike simple impedance, was found to be directly relatable to defined culture characteristics such as cell coverage (Wegener et al., 2000). A model using complex

impedance measurements as inputs can also be used to derive additional culture information such as the barrier resistance or average capacitance of the plasma membrane (Giaever, 1991). The cell behaviour captured by those measurements will be different depending on the frequency of the alternative current applied (Bagnaninchi and Drummond, 2011; Giaever and Keese, 1993). Low frequencies inform on the space between the cells (the current flows mostly in-between cells) while high frequencies are sensitive to cell coverage (the current can flow directly through the cell membrane).

ECIS has been extensively used to monitor various cell behaviours, including attachment, motion and morphological changes. The system is compact and can be kept in a conventional incubator, enabling continuous, real-time data acquisition. It requires the use of specialised and costly culture vessels with integrated gold electrode arrays. The arrays can represent a sizable fraction of the culture area, which could lead to undesired behaviours due to the cells attaching to and interacting with a non-standard culture substrate. The measurements obtained for cells growing on the gold electrodes might thus not be representative of the rest of the culture (e.g. cells growing on tissue culture polystyrene) and do not necessarily generalise to conventional culture vessels that do not contain embedded sensors. The last drawback concerns the interpretation of the data. As the impedance measurements are sensitive to a wide range of cell behaviours, it can be difficult to isolate the relevant information. Data interpretation therefore often requires the use of models (Chen et al., 2012; Giaever, 1991). This may require additional calibration efforts to obtain the correct application-specific parameters for the model (e.g. new cell line or different cell behaviour studied).

1.2.7 Indirect monitoring

It is often possible to derive useful information about cells in culture by monitoring their environment. For example, analysis of the culture medium can inform on the metabolism of the cells (e.g. consumption of nutrients or production of waste metabolites). Spent medium can be easily sampled during a culture (if a small volume is required) or at the time of a scheduled medium exchange. Commercial instruments (such as the BioProfile line from Nova Biomedical) are equipped with amperometric electrodes with embedded immobilised enzymes for the detection of glucose, glutamine, lactate and glutamate (Bawn et al., 2013). Amino acids can also be detected using chromatographic methods (Hanko and Rohrer, 2004; Perry et al., 1968).

Instead of relying on intermittent sampling of the culture medium, some of those molecules can be measured in real-time. Screen-printed sensors coated with immobilized enzymes for the online monitoring of glucose, glutamate and lactate were recently described (Eibl and Eibl, 2011). Similarly, optical probes based on fluorescence lifetime measurements can be used to non-invasively monitor pH and dissolved oxygen (Hanson et al., 2007; Naciri et al., 2008).

1.3 Automated processing and analysis of light microscopy images

The previous section showed the wide palette of methods available to cell culture researchers desiring to quantify various aspects of live adherent cell cultures. The choice of an optimal method for a specific application is a balancing act, where multiple criteria have to be taken into account. For long-term culture monitoring, it is important to ensure that the chosen method results in as little disruptions to the culture as possible. Likewise, compatibility with conventional culture ware and protocols is required so that experimental workflows do not have to be significantly altered to accommodate the monitoring method. Light microscopy fits these requirements remarkably well. Indeed, light microscopy methods in general, and phase contrast microscopy (PCM) in particular, are widely recognised as the *de facto* techniques for the inspection of live adherent cell cultures. In many cases, however, PCM images are relegated to documentation and archival purposes.

Recently, there have been efforts to apply image processing and machine vision techniques to the analysis of PCM images with the goal of turning them into a source of reliable and quantitative data for the characterisation of live adherent cell cultures. This is not without challenges, as unlike high contrast images such as those acquired using immunofluorescence microscopy, PCM images are often very noisy due to the presence of artefacts, including bright halos around cellular objects and the shade-off effect (see section 1.2.2). In this section, applications made possible by processing and analysis of PCM images will first be briefly described. A more technical overview of existing approaches, including specialised algorithms and generic approaches taking advantage of recent developments in machine learning, will then be presented. Finally, some key considerations for the application of these methods in real-world experimental scenarios (e.g. non-optimal imaging conditions in a cell culture laboratory) will be outlined.

1.3.1 Applications in adherent cell culture monitoring

The different types of analyses that can be carried out using an image processing algorithm are not necessarily broader in scope than what a human expert can achieve given an infinite amount of time. As a rule of thumb, if a human expert cannot perform a vision task (e.g. counting cells in a microscopy image) after sufficient training then there is little chance that a machine vision algorithm would be able to. Instead, the application of image processing methods to microscopy images can be thought of having an operator who can tirelessly, consistently, and very rapidly annotate and analyse images while remembering all previously seen images to potentially find correlations between them (e.g. frame of a time-lapse movie). Image processing is thus most useful to automate measurements carried out based on microscopy images and to minimise their subjectivity and variance.

Any stage of a typical adherent cell culture experiment where cells are visually inspected (see Figure 1-1) could in principle be improved using automated image processing and analysis. A simple yet often neglected application is the documentation of routine cell culture maintenance steps. Frequent imaging and characterisation of cell cultures would allow the early detection of deviations from expected growth patterns, for example due to changes in cell phenotype as a result of unsuitable handling (Veraitch et al., 2008). The same measurements could also be used for decision-making during experiments, whereby the timing of a given procedure (e.g. addition of a molecule to the culture medium) would be based on quantitative data rather than on an arbitrary and rigid schedule that remains the same regardless of the state of the cells.

For documentation and decision-making purposes, the focus is usually put on measurements of the cell population size. These include enumeration of individual cells, from which a measure of cell density can be derived (i.e. number of cell per unit area), and determination of culture confluency (i.e. fraction of the growth area occupied by cells). Likewise, other measurements such as morphology (i.e. cell shape) could be of interest in many instances, for example for the determination of cell phenotype. These measurements are already part of most cell culture protocols, albeit in a qualitative form based on visual inspection of the cells. The application of machine vision techniques to microscopy images can provide a standardised and robust framework for measurements of cell culture characteristics in datasets ranging from a few images (e.g. documentation of routine procedures) to millions of files generated using high throughput imaging systems (see sections 1.1.4 and 1.2.2).

There is also a need to match the emergence of experimental setups that can accommodate long-term imaging of adherent cell cultures (see section 1.2.2) with adequate analysis capabilities so that dynamic cell behaviours can be quantified. A typical example is the tracking of cell migration across frames of time-lapse movies and subsequent quantification of their speed and persistence. This information is critical to gain a thorough understanding of many biological processes, including wound healing (Cai et al., 2007).

The next section will introduce image processing and machine vision methods that make these applications possible. Two broad categories of algorithms will be discussed. Firstly, PCM image segmentation algorithms, which are the main focus of this work, will be described (section 1.3.2). These methods allow the detection of the boundaries of cellular objects in PCM images and their output can then be used as a basis for further analysis such as area or shape measurements. In contrast, object detection algorithms are used for the determination of the approximate location of individual cells in PCM images (section 1.3.3). Object detection methods are mainly used for the enumeration of cells as well as their tracking across frames of time-lapse movies.

1.3.2 Phase contrast microscopy image segmentation

Typically, segmentation of a microscopy image consists in classifying pixels as either foreground (i.e. cells) or background (Gonzalez and Woods, 2007). The output of segmentation is usually a binary image (or binary mask) of the same dimensions as the input image where values of zero and one indicate that the corresponding pixels in the input image were classified as background and foreground (i.e. cells), respectively (Figure 1-15). Binary images are usually obtained by thresholding, whereby pixel values above a set threshold will be given a value of one in the binary mask, and zero otherwise (see section 2.1.2.1 for a detailed explanation of thresholding techniques). This binary image is then used as a basis for further analysis, for example for the determination of the fraction of an image occupied by cells.

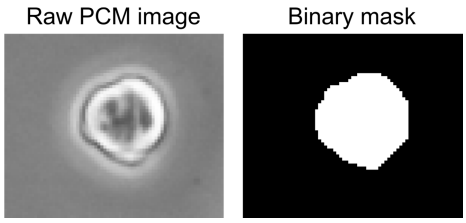


Figure 1-15 Example of a PCM image and the corresponding binary segmentation output. By convention, background pixels are represented in black and foreground pixels in white.

Segmentation of PCM image is particularly challenging due to artefacts that are intrinsic to the method: large, bright halo artefacts around cellular objects and the lack of contrast between the cells and the background due to the shade-off effect (see section 1.2.2 for details). Segmentation algorithms thus have to deal with those issues in order to produce satisfactory results. A selection of algorithms that were previously reported in the literature is discussed below and summarised in Table 1-2. The techniques presented here include contrast filters, active contour models, watershed transform, image formation models, and learning-based approaches.

Table 1-2 Selection of publications describing methods for the segmentation of cells on phase contrast microscopy images. References were chosen according to their relevance to the work describe in this thesis. The processing times shown here are based on the information provided in the corresponding paper and are for illustration only as they were obtained using implementation coded in different programming languages and running on different machines.

Ref	Method	Processing time[1]	Sensitive to halo artefacts	Application
(Debeir et al., 2008)	Ensemble of weak watershed classifiers	?	?	Segmentation of individual cells
(Bradhurst et al., 2008)	Contrast filter followed by halo correction using dynamic programming	?	No	Classification of cells based on phenotype
(Kazmar et al., 2010)	Learning of texture and intensity features	?	No	Segmentation and tracking of individual cells
(Yin et al., 2010).	Bag of local Bayesian classifiers	~50s	No	Segmentation of individual cells
(Theriault et al., 2011)	Edge and contrast filters	?	Yes	Classification of cells based on phenotype
(Topman et al., 2011).	Multi-scale contrast filter	< 1s	Yes	Confluency determination
(Seroussi et al., 2012)	Parameterised active contour models (GVF, DGVF)	~8.6s	No	Tracking of single cells in sub-confluency images
(Ambühl et al., 2012)	Geometric active contour model	~500s	No	Tracking of single cells in sub-confluency images
(Yin et al., 2012)	Image formation model for restoration of artefact-free images	~2s	No	Tracking of single cells
(Juneau et al., 2013)	Contrast filter	?	Yes	Confluency determination and culture monitoring
(Su et al., 2013b)	Classification of "phase homogeneous atoms" using graph cuts	?	No	Segmentation of cells in a sequence (e.g. wound healing)

1. For a 1280×960 pixels image. Estimated based on information provided in the corresponding manuscript

1.3.2.1 Contrast filters

Regions of PCM images that correspond to cells are usually associated with a high variability in pixel intensity values due to the presence of cellular structures such as the cell membrane, cytoplasm or nucleus. In comparison, the intensity in the background regions is relatively homogeneous. In other words, background pixels are likely to be surrounded by pixels of similar intensity values, while intensity values in the close neighbourhood of cell pixels are likely to be vastly different. Application of contrast filters such as local standard deviation and variance filters produces an image whose pixel values are proportional to the intensity variability in the neighbourhood of the corresponding pixel in the input image, thus allowing distinguishing between cell and background regions (Figure 1-16). The scale of a contrast filter is the size of the local neighbourhood (i.e. window) considered around each pixel. Typically, the scale is chosen so that it is large enough not to be overly skewed by uninformative noise while

remaining sufficiently small to be representative of local rather than image-wide variations in intensity. Alternatively, multiple filter scales can be combined. Contrast filters, or variations thereof, have been widely used for the segmentation of PCM images.

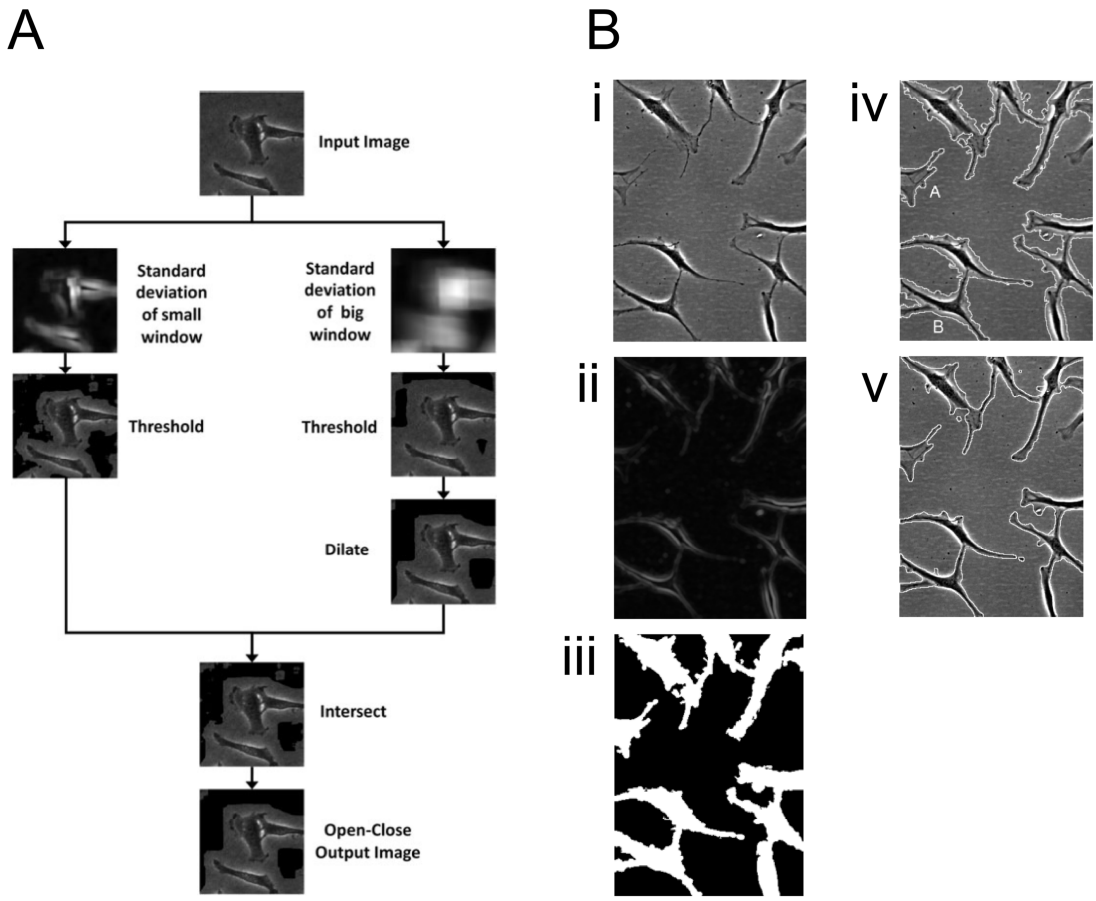


Figure 1-16 Examples of contrast filter approaches for PCM image segmentation. (A) Segmentation of PCM images based on the application of standard deviation filters at two scales, followed by thresholding and morphological operations. Figure adapted from (Topman et al., 2011). (B) Refined PCM image segmentation approach that includes correction of halo artefacts. The PCM image (i) is first filtered using a standard deviation filter (ii), before thresholding the resulting image to produce a binary mask where cell regions are correctly detected, but where halo pixels are misclassified as cells (iii). A post-processing algorithm based on image intensity gradient is then used to produce a refined segmentation that corrected for halo artefacts. Figure adapted from (Bradhurst et al., 2008) © 2008 IEEE.

A method that combined edge detection with local standard deviation filters for the segmentation of cells on PCM images was described (Theriault et al., 2011). First, edge images at three different scales were computed and thresholded. The local standard deviation transform of the image was then efficiently computed using the integral image approach (Porikli, 2005) and subsequently thresholded. The final segmentation mask was a combination of the binary masks obtained from the edge and local standard deviation images. Similarly, another study proposed a method based on the computation of local standard deviation images at two scales (Topman et al., 2011). The binary masks obtained from the standard deviation images after thresholding were then combined to form the final segmentation mask (Figure 1-16.A). Another approach was similar but used a local range filter to detect regions of high pixel intensity variability instead (Juneau et al., 2013). The processing time for these methods was low as the computation of contrast filters (local standard deviation and range filter in these cases) is based

on highly efficient and optimised algorithms. A downside of the naïve use of contrast filters is that halo artefacts are in most cases misclassified as cell pixels.

The issues related to the misdetection of halo artefacts as cellular objects were alleviated by combining contrast filtering with dynamic programming (Bradhurst et al., 2008). First, a local standard deviation image was computed and thresholded (Figure 1-16.B). Halo artefacts in the resulting segmentation mask were then corrected by scanning the mask in four directions (i.e. top to bottom, bottom to top, left to right, and right to left) and detecting sudden changes in the intensity gradient sign, which were usually associated with the transition from the bright halo to the dark cell border. By correcting for the halo artefacts, the accurate contours of the cells were detected in most cases.

Methods based on contrast filter are highly efficient and when combined with a halo artefact correction step, could potentially produce high quality segmentation results despite their apparent simplicity. They are therefore good fits for applications that only require labelling of pixels as 'cell' or 'background' rather than relying on the detection of cells as individual objects.

1.3.2.2 Active contour models

Active contour models (also commonly referred to as 'snakes') are energy-minimising parametric curves (often splines) that can be used to detect edges and boundaries on images (Kass et al., 1988). The energy function to minimise has two main terms: the model's internal energy and image forces. The internal energy depends solely on properties of the spline itself (e.g. length, curvature). As the spline is usually expected to converge towards the contour of the object, it needs to shrink while maintaining its rigidity. In order to do so, the internal energy is often based on the elasticity and bending energies. External forces define what features of an image will attract the spline. A typical example is the use of the image gradient as an external force so that the model would be attracted to contours with high intensity gradients (i.e. edges). The position of the snake is usually initialised in the close neighbourhood of the object to be delineated based on preliminary segmentation using any other method that can consistently estimate the location of the object. The cost-function is then iteratively minimised until the snake conforms to the contour of the object. This type of model is known as parameterised active models. A common issue with this approach is that it is very sensitive to initialisation. The changes in image gradient caused by the contour of an object might only be local and if the spline is initialised outside of this range, it won't be attracted to the contour of the object.

The issues can be partially alleviated by using gradient vector flow (GVF) as the snake's external force (Xu and Prince, 1998). GVF is computed as the diffusion across the image of the aforementioned image gradient vectors. In addition to allowing snakes to converge even when

not initialised in close proximity of the object, it also greatly improves the detection of concave contours (Figure 1-17.A). GVF snakes were successfully applied to the segmentation of well-separated cells in PCM images (Seroussi et al., 2012). The snake was initialised by roughly segmenting the PCM image based on edge detection followed by thresholding. Using conventional GVF, the snake converged to the halo artefact surrounding the cells (Figure 1-17.B.i). The first snake is then used as the initialisation point for a second model, this one driven by a directional GVF (Schoepflin, 2001) in order to determine the actual cell boundary (Figure 1-17.B.ii). The segmentation result was used to track cells across multiple frames of a time-lapse movie using a previously described correlation-based approach (Wilson and Theriot, 2006).

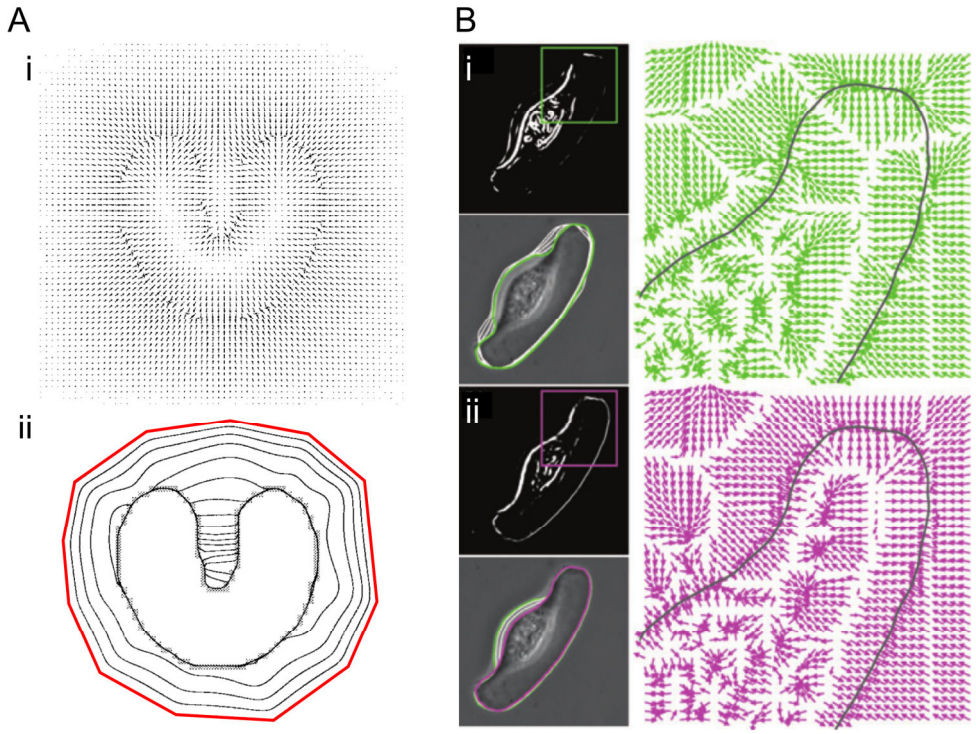


Figure 1-17 Active contour models for PCM image segmentation. (A) Synthetic example of a gradient vector flow (GVF) snake. (i) External GVF forces guiding the convergence of the snake represented by arrows. (ii) Iterative convergence of the snake with the initial state shown in red. The snake converges until it conforms to the U-shaped object despite its concave nature. Figure adapted from (Xu and Prince, 1998) © 1998 IEEE. (B) Application of GVF snakes to the segmentation of cells in PCM images. (i) A first snake converges based on GVF calculated from the PCM image's edge map, shown in the top left insert. The final contour of the snake is shown overlaid on the original PCM image in the bottom-left insert (ii) Second snake used to refine the contour detection based on directional GVF (DGVF). DGVF was computed from a directional edge map (top-left insert) where all vectors have an outward component relative to the cell boundary as approximated by the contour of the first snake (green outline). The final contour is shown as a purple outline overlaid on the original PCM image in the bottom-left insert. Figure adapted from (Seroussi et al., 2012), reproduced with permission from the publisher (Wiley).

The approaches discussed above define the contour as a collection of control points that give a curve after interpolation, and are thus explicit (also referred to as parametric). In contrast, alternative models have been proposed where the contour is implicitly defined as the zero level-set of a continuous function (Li et al., 2005). The main advantage being that the contour can naturally split and merge as necessary during convergence. An implicit contour approach was taken for the detection of single cells on high resolution PCM images (Ambühl et al., 2012).

Firstly, a rough segmentation was obtained by automatic thresholding of the gradient image. The resulting binary mask was then used for the initialisation of a first geometric snake on the original grayscale image. This first snake converged towards the bright halo artefact surrounding the cell. This contour is then refined by switching to a top-hat transform of the original grayscale image, which removes bright features such as the halo artefact. After convergence of the snake on the top-hat transform image, the last step is carried out on the original grayscale image, where a final snake will converge towards the actual cell contour. The algorithm was thoroughly validated (e.g. effect of image noise on the algorithm performance) and enabled the analysis of minute dynamic cell shape changes as well as the creation of high-resolution maps of cell behaviours such as protrusion and retraction.

The implementation of both explicit and implicit active contour models enabled the detection of cell contours in PCM images. A down side of these approaches, however, is their relatively high computational complexity, which results in long processing times for high-resolution images. Active contour methods are thus best suited for applications where highly accurate detection of cell contours is critical (e.g. characterisation of subtle cell membrane movements in high resolution time-lapse microscopy movies). For other uses, alternatives methods such as contrast filters might be better fits if it is acceptable to sacrifice the accuracy of the detected cell contours in favour of much shorter and practical processing times.

1.3.2.3 Watershed transform

The watershed transform has been widely used in many image processing applications (Roerdink and Meijster, 2000). If an image is interpreted as a topographic relief, pixels of high intensity values can be related to hills and those of low values to valleys (Figure 1-18.A). In the case of the gradient image, the hills will correspond to object edges while other regions (i.e. interior of objects and background) will correspond to valleys. The watershed transform can be thought as having water rising from local minima of the gradient image with the constraint that water from different sources (i.e. different local minima) cannot merge. The watershed transform results in the partition of the image in catchment basins and watershed lines. The latter correspond to the segmentation boundaries. Segmentation using a watershed transform frequently leads to over-segmentation due to the presence of a large number of local minima. The quality of the segmentation can be drastically improved by having the water rising from pre-defined markers instead, a technique called marker-controlled watershed transform (Debeir et al., 2008).

The segmentation result will be highly dependent on both the location of the markers and the image noise levels, which makes the technique normally difficult to apply to PCM images. These issues were alleviated by introducing randomness in the marker location and in the image intensity levels (Debeir et al., 2008). The approach was somewhat relatable to bootstrap

aggregating methods in machine learning where random selection of training data and the creation of ensemble of weak learners lead to improved classification performance (Breiman, 1996; Freund and Schapire, 1997). Noise was added to the original image by addition of randomly oriented intensity slopes (Figure 1-18.B). Similarly, the locations of the cell markers were altered based on random noise. This was repeated multiple times and a counter kept track of how many time each pixel of the image was part of the random cell basins generated after application of the watershed transform. This yielded a probability map (consensus of the random watershed transforms) that could then be thresholded to obtain the contour of the cells. This approach was promising but still relied on the prior knowledge of marker locations, and thus was unsuitable for automated cell monitoring.

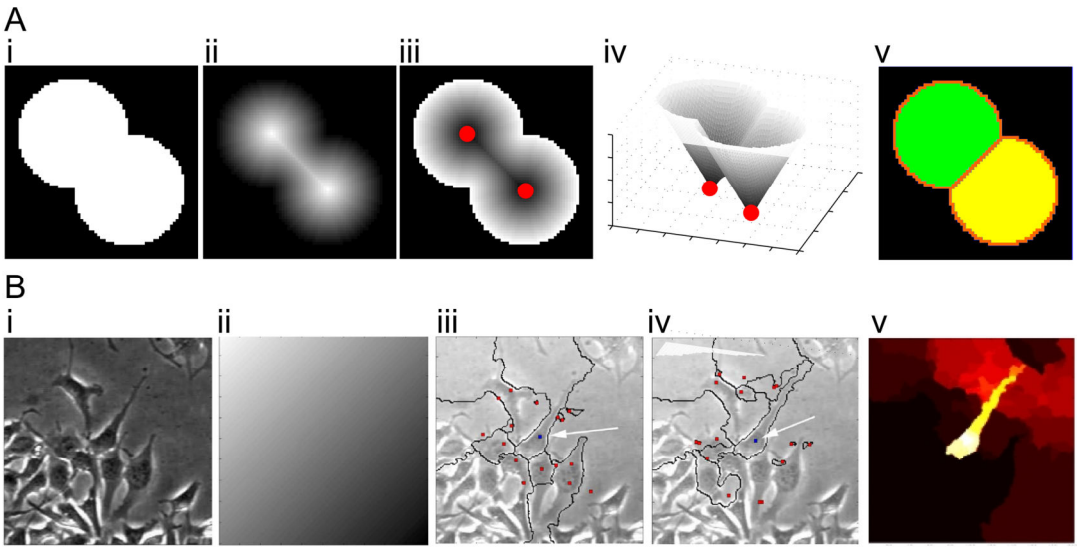


Figure 1-18 Example of application of watershed transforms to image segmentation. (A) Synthetic example demonstrating the principle behind watershed-based segmentation. (i) Image containing touching binary circles that should be separated during segmentation. The distance transform is first computed (ii) and inverted (iii), so that the centres of the circles (shown with red markers) correspond to local minima. (iv) 3D representation of (iii), highlighting the "catchment basin" topology of the inverted distance map. (v) Result of the application of the watershed transform. The two circles (shown here in green and yellow) are separated as the water rising from separate local minima cannot merge. The orange line corresponds to watershed lines. (B) Application of the watershed transform to PCM image segmentation. (i) The PCM image to segment. (ii) Random intensity slopes are generated and added to the image. (iii) The cell to segment (shown by a white arrow) is manually marked by the user (blue dot). Random background seeds are added (red dots). Seeded watershed segmentation is then carried out similarly to what was shown in (A). (iv) The process is repeated after addition of another random intensity slope, producing slightly different results. (v) After repeating the process multiple times, the contour of the cell is given by the consensus segmentation across the different iterations. Bottom panel adapted from (Debeir et al., 2008), © 2008 IEEE.

1.3.2.4 Modelling image formation mechanisms

By exploiting the knowledge of the physical processes leading to the formation of a PCM image (see section 1.2.2), it was suggested that images could be restored so that halo artefacts would effectively be corrected and the contrast between cellular regions and background greatly enhanced (Yin et al., 2012). The resulting artefact-free image can then be segmented using a simple intensity threshold. A potential issue was the necessity to obtain dimensions related to the optics used in the microscope, which are typically not disclosed by the manufacturers. A study that applied this algorithm for online monitoring of adherent cultures reported

underwhelming segmentation performance when compared to ground truth data generated by a human expert (Ker et al., 2011). While very elegant and ingenious, this image formation model had yet to demonstrably produce segmentation performance superior to the alternative approaches described here-in.

The image formation model did not account for variations in cell physical properties (e.g. thickness) during the course of a culture (e.g. due to mitosis, differentiation or apoptosis), which in turn impacts the image formation mechanism. When images contained a heterogeneous population of cells (i.e. possessing different physical properties), the assumptions made by the model were no longer valid and the reconstructed artefact-free image was therefore incorrect. This phenomenon explains the relatively poor segmentation performance reported when the model was applied to long-term culture monitoring. These issues were partially alleviated by a subsequent study that segmented PCM images based on local diffraction patterns (Su et al., 2013b). A dictionary of diffraction patterns was created based on the findings from the aforementioned image formation model (Su et al., 2013a; Yin et al., 2012). The idea was that a PCM image could be represented as a linear combination of these diffraction patterns. Pixels with similar diffraction patterns were then clustered together in 'phase-homogeneous atoms' using a method related to superpixels (Achanta et al., 2012). The atoms were then classified as dark cells, bright cells, background, and boundary regions. Classification was carried out by partitioning a graph of similarities between atoms using graph cuts. The authors reported significant improvements in segmentation performance over that obtained using the original formulation of the image formation model. A potential caveat was that the optimal results were obtained using a semi-supervised approach, where a number of atoms had to be manually labelled by the user. This would thus preclude the use of this algorithm for online monitoring.

1.3.2.5 Learning-based or trainable segmentation

Recently, there has been an effort to develop generic approaches for machine vision problems that would not necessarily require any *a-priori* knowledge about the images to process but instead relied on supervised machine learning techniques to devise the most optimal mapping between image features (e.g. intensity, texture) and the desired output. Such techniques are termed learning-based or trainable segmentation in this work to differentiate them from most other methods described above, which relied on the knowledge of specific features of PCM images in order to develop specialised segmentation algorithms.

A study used machine learning to partition an image into multiple cellular features (e.g. nucleus, membrane, halo) based on texture features (Kazmar et al., 2010). Features were computed for each 5×5 pixels neighbourhoods and included texture features first introduced by Haralick (Haralick, 1979) as well as the intensity value of the central pixel. Various machine-learning

classifiers were assessed, including support vector machine (SVM), decision trees and random forest. The latter produced the best segmentation results.

Alternatively, PCM images were segmented based on a bag of local Bayesian classifiers (Yin et al., 2010). The core principle behind this approach was the generation of an ensemble of classifiers (here termed a bag) that were experts in classifying a certain type of image patterns (e.g. dense or sparse cell regions). For classifier training, histograms of intensity values for about 5000 local windows centred at different location in an image were constructed. The pair-wise distance between these histograms was computed and the resulting similarity matrix clustered using spectral grouping (Fowlkes et al., 2004). A Bayesian classifier was trained for each cluster, making it an expert to deal with the corresponding type of intensity histogram. When processing a new image, local intensity histograms were first computed for each pixel of the image. These histograms were then compared to the average histogram for each clusters obtained during training to decide which Bayesian classifier is most appropriate for the classification of a particular pixel. Processing time was approximately 50 seconds for a typical PCM image, although the authors noted that the pixel-wise segmentation approach taken would enable the use of parallel computing to significantly decrease processing time. Impressive results were obtained for the segmentation of phase contrast time-lapse sequences using this approach.

Learning-based (trainable) segmentation is still in its infancy, in particular its application to PCM images, but already achieves performance that rivals that of state-of-the-art specialised algorithms. Software package such as “Ilastik” (Sommer et al., 2011a) and “FIJI’s trainable WEKA segmentation plugin” (Hall et al., 2009; Schneider et al., 2012) are becoming increasingly popular for the segmentation of biomedical images. Both are based on a similar pixel-wise segmentation scheme using a random forest classifier (Breiman, 2001a). Each pixel of an image is classified into one of several pixel classes (background or foreground for binary classification) based on image features. Both “Ilastik” and “trainable WEKA segmentation” allow the combination of a wide variety of image features, including intensity, texture, and edge maps. Random forest was used as it can handle the classification of millions of pixels while maintaining low memory usage as well as fast training and classification times. The efficiency of random forest also enables interactive segmentation applications whereby the classifier is updated almost in real-time as the user annotates the image. Visual feedback then indicates ambiguous regions that require additional input from the user to improve segmentation. Interactive segmentation thus removes the need for tedious parameter optimisation phases that are normally associated with image processing algorithms, albeit at the risk of impacting on the generalization of the approach to other datasets.

This work will investigate whether such generic trainable segmentation algorithms are suitable for imaging-based monitoring of adherent cell cultures and how they compare to specialised algorithms.

1.3.3 Object detection and tracking on phase contrast microscopy images

The segmentation approaches described above can be used for the detection of object in PCM images. However, simpler and more efficient methods exist for applications that do not require an accurate contour of the cells to be detected. An example of such an application is the enumeration of adherent cells on PCM images. If the morphological information is not of interest, it is sufficient to detect individual objects so they can be counted, for example based on a template matching method (Usaj et al., 2011). Images were convolved with a Laplacian-of-Gaussian (LoG) cell model. At each location, a correlation factor measured how closely local image structures matched the LoG model. Cell centres were then assumed to correspond to local correlation factor maxima. The three detection parameters (cell diameter, minimum distance between cells, and threshold) were automatically determined from the training data using an artificial neural network. An average counting error of ~10%, which increased with the number of cells in the image, was reported.

Object detection is also frequently used to track the positions of cells in migration studies. Cells were tracked across frame of a time-lapse microscopy movie using optical flow (OF) registration (Hand et al., 2009). Registration is the process of mapping objects across multiple images so that their motion can be measured. OF is a widely used set of methods for motion estimation based on the computation of the flow of pixel intensities across two images. OF has the big advantage of enabling the registration of objects across frames of a movie despite changes in shape, which makes it particularly suitable for the tracking of cells that often have very dynamic morphologies. After an initial rough segmentation step, cells were tracked on subsequent frames automatically. A similar approach using an alternative formulation of fluid registration (Kuska et al., 2008) was employed to track mouse embryonic stem cell colonies over time (Scherf et al., 2012). In addition to motion estimation, this particular implementation of fluid registration enabled the tracking of merging and splitting events, which are of significance for colony forming cell lines.

In a different approach, the mean-shift algorithm was adapted to enable the tracking of cells with changing morphologies (Debeir et al., 2005). The mean shift algorithm is an iterative process to determine the local mode of a distribution (Cheng, 1995). In the context of an image, the mean-shift algorithm will converge towards the centre of a given intensity configuration (e.g. dark pixels on bright background). For application to microscopy images, the original approach was modified to also detect mitotic events so that cell lineages could be automatically

tracked. The proposed mean-shift algorithm enabled tracking over long periods (>3 days), with the downside of relying on a manual initialisation step by the user.

Images acquired out-of-focus can carry information not necessarily present in in-focus images. This information can potentially aid the segmentation process. A typical application of this phenomenon is the detection of cells in bright field microscopy images where stacks of out-of-focus images with increased contrast enabled the detection of cells that are otherwise almost entirely invisible in in-focus images (Dehlinger et al., 2013; Selinummi et al., 2009). A major downside of such these approaches, however, is the need to acquire multiple images per field of view, thus increasing imaging time and requiring the use of a microscope equipped with a motorized stage. A method proposed for the detection of cells in confluent epithelial monolayers alleviated these issues (Flight et al., 2014). The out-of-focus phenomenon was simulated by filtering PCM images with two mean filters, one of radius smaller than the average cell and the other of radius larger than the average cell. Subtraction of the two mean images resulted in a new image where boundary and cell regions corresponded to low and high pixel values, respectively. Application of a simple threshold allowed segmentation of individual objects, which corresponded to the interior of cells. While the segmentation output did not represent the cell boundaries, and thus could not be used for shape analysis or area determination, it enabled their enumeration.

Some of the aforementioned object detection methods were combined with a machine learning framework in order to improve cell detection performance on PCM images (Pan et al., 2009). First, a pre-processing step detected pixels that are very likely to be background so they can be discarded from the analysis. This was done by determining the fluctuation energy of each pixel, computed from the response to a series of Laplacian filters. Background pixels were determined by thresholding the fluctuation energy image. Pixels that corresponded to local intensity extrema were considered as candidate cell pixels. Using the previously mentioned mean-shift algorithm, the location of these candidates was further refined. To determine whether the remaining pixels were cell or background, a support vector machine (SVM) was used to learn local features such as mean intensity, mean gradient magnitude, and histogram of oriented gradients. The output of the SVM corresponded to the probability of a given pixel to be a cell pixel given the input features. In order to group together the remaining cell pixels if they are within the same cells, another SVM was trained using pairwise features based on the optimal path between the two candidates. If two points were within the same cell, a path could connect them without going over any bright ridges, which were usually associated with the background of the image. The output of this SVM was the probability of the two points to be in the same cell given the pairwise features. The combination of all these methods allowed reaching very high object detection performance.

1.3.4 Considerations for real-world applications

There are significant hurdles to overcome in order to successfully transition from a theoretical proof of concept algorithm to a tool that can be integrated into adherent cell culture workflows. A major aspect is the performance of the method in real-world scenarios. Indeed, while good practices have been established for the evaluation of microscopy image processing algorithms, some aspects remain challenging to thoroughly characterise and are not necessarily considered in theoretical investigations (Jannin et al., 2006; Jannin, 2012). These aspects include but are not limited to generalisation, processing time, and accessibility.

Generalisation of an algorithm is particularly important due to the inherent variable and unpredictable nature of microscopy data. Even for a given cell type, visual features can vary greatly depending on the experimental conditions (e.g. culture vessel used, seeding cell density). The imaging setup (e.g. microscope model, camera specifications) and protocols are also likely to differ significantly between laboratories. What's more, image acquisition during experimental investigations is rarely carried out in optimal and controlled conditions, potentially resulting in image quality degradation. Depending on the scope of the proposed method, it is therefore necessary to carefully devise an adequate validation scheme in order to evaluate sensitivity to these different sources of variability. Failure to do so might lead to an overestimation of the algorithm performance in relevant experimental scenarios and could thus significantly hinder its applicability.

The type of application will also dictate the requirements in terms of processing time. For instance, a tool that is to be used routinely as part of adherent cell culture workflows for decision-making or documentation needs to be quick enough so that it does not incur significant additional analysis time when compared to visual inspections. As such, the processing time should not be significantly higher than the time required to acquire the images themselves. For more advanced applications such as time-lapse analysis (e.g. cell tracking), the image processing and analysis are usually carried out offline after completion of the experiment. In those cases, maintaining a low processing time is not as critical. The balance between absolute performance and processing time should therefore be chosen on a per-application basis.

Another key consideration when developing an image processing tool to be used in research laboratories is its accessibility. It is unlikely that a user would spend a significant amount of time tweaking parameters in order to get the algorithm to produce satisfactory results. This can be alleviated by devising methods that generalise well and have low parameter sensitivity. As an alternative, the complexity can be hidden by facilitating optimal parameter values estimation in a user-friendly and automated manner.

1.4 Towards a high performance and scale-independent monitoring methods for live adherent cells

The preceding sections introduced the many concepts underlying this fundamentally inter-disciplinary research project. The overarching theme was the transition of cell biology towards a systematic and quantitative experimental framework that is suitable for the description and documentation of biological processes of ever increasing complexity. This transition is occurring concurrently for all aspects of cell culture, from the design of the culture vessels to the methods used for the generation of experimental data and its subsequent analysis. Often, it is done by borrowing from well understood and characterised methodologies developed in other fields, for example in engineering, material science, and computer science.

The limitations of conventional cell culture approaches were also witnessed first-hand from personal experience in several cell biology laboratories, which was corroborated by extensive discussions with fellow researchers. The lack of standardised methods for the non-invasive and rapid quantification of adherent cell culture is a major hurdle as most results tend to be communicated and reported in a very qualitative way unless destructive end-point methods are used. This is an obstacle for the establishment of robust cell culture protocols where decision-making should be based on reliable data. Mundane, yet necessary, fundamental tasks such as the establishment of a growth curve remain surprisingly difficult to perform. Likewise, conventional culture vessels whose design and mode of operations have not significantly changed in more than a decade are not suitable for the investigation of complex processes. A typical example is the self-renewal and differentiation of pluripotent stem cells. *In-vivo*, this process relies on a large number of cues that are modulated overtime. These very dynamic conditions cannot be replicated in a laboratory by growing cells in vessels left unmonitored in incubators, where the control over the culture conditions is limited to infrequent exchanges of the culture medium and sporadic adjustments of incubator settings.

Similar issues potentially impact all areas of biology where adherent cells play an important role. This worked aimed at addressing them in two steps. Firstly, develop and thoroughly validate automated processing and analysis methods for images acquired using phase contrast microscopes, which are already in place in most cell culture laboratories. Secondly, combine the best approach devised with a novel microfabricated bioreactor to allow online monitoring of cell culture in a highly controlled environment. In order to achieve this, the monitoring method should thus be scale-independent (in reference to the scale of the culture vessel and not that of the cells).

In order to achieve this aim, algorithms for the segmentation of cells on phase contrast microscopy (PCM) images (i.e. labelling of each pixel as either cell or background) had to be developed. Two fundamentally different approaches were taken to account for different type of

images and applications. The first one favoured low processing times and simplicity in order to offer an alternative to qualitative visual inspection for standard cell culture workflows while the second accommodated complex experimental scenarios such as co-cultures and those where visual features of cells are expected to change drastically during the course of a culture.

The first method relied on a specialised algorithm whose design and implementation was informed by the properties of PCM images (Chapter 2). It was based on computationally inexpensive linear contrast filters to detect regions of images likely to contain cells based on high local variations in pixel intensity values. Segmentation was further improved using a *post-hoc* dynamic programming algorithm that corrected for halo artefacts that are intrinsic to PCM images. The approach was thoroughly validated and compared against previously described PCM image segmentation methods. More importantly, its sensitivity to variations in imaging conditions (i.e. microscope model, camera type, and illumination intensity) was also assessed in order to better understand how it would fare in real-world situations.

This first segmentation algorithm was used as a basis for the determination of culture confluency (i.e. fraction of the culture area occupied by cells), the estimation of culture density (i.e. number of cells per unit area), and the analysis of morphological attributes as well as for the interpretation of fluorescence microscopy data (Chapter 3). Application of culture monitoring based on those measurements was demonstrated for a number of experimental scenarios involving mouse embryonic stem cells. The palette of cell responses investigated included cell proliferation, cell death, growth arrest, transient morphological changes, and changes in gene expression patterns.

Because the approach based on contrast filters relied on *a priori* knowledge about the structure and properties of PCM images acquired during conventional adherent cell culture experiments, it was not suitable for the processing of images that significantly deviated from what was expected in such settings. An alternative approach based on a generic machine learning framework was thus investigated in an attempt to alleviate this limitation (Chapter 4). Using a technique termed trainable segmentation, a random forest classifier segmented PCM images by labelling each pixel in one of two classes (e.g. ‘cell’ and ‘background’) based on local image features such as intensity and texture. Segmentation performance was compared with that obtained with the specialised algorithm for standard cells versus background segmentation tasks. The trainable segmentation approach was then evaluated for more complex segmentation scenarios involving the discrimination between two cell types.

The inherent ability of the trainable segmentation algorithm to learn how to detect a wide range of cell appearances made it the perfect fit for the monitoring of long-term cultures in a microfabricated culture device during which phenotypes, not usually observed in shorter term cultures, emerged. Indeed, by continuously perfusing fresh culture medium, the cells were

exposed to a well-controlled and stable soluble microenvironment and could be cultured without detachment beyond what is normally achieved in flasks or well plates. A custom-developed, fully automated image acquisition setup combined with the trainable segmentation algorithm enabled online monitoring of the entire culture area and allowed, for example, the detection and analysis of local growth patterns (Chapter 5). The monitoring of both mouse and human embryonic stem cell cultures in the microfabricated bioreactor was demonstrated.

The main novel contributions of this work are the following:

- Novel PCM image segmentation algorithm that outperforms similar methods found in the literature for the set of images considered.

- A new trainable segmentation scheme based on multi-scale local image feature histogram encoding for important spatial information and context. This scheme outperformed state-of-the-art software packages for the segmentation of complex images of co-cultured Human embryonic stem cells.

- Monitoring of spatial and temporal growth patterns in a microfabricated bioreactor using the aforementioned trainable segmentation scheme, providing information that was previously unattainable in such systems.

Chapter 2

Segmentation of phase contrast images using contrast filters

Observation of cells using light microscopy is a de facto standard in cell culture. Visual inspection is quick, non-disruptive, and enables an experienced operator to rapidly assess various culture characteristics. However, these observations remain qualitative and subjective. Other non-invasive methods are often limiting due to high cost, incompatibility with conventional culture ware, or potential adverse effects on cell phenotype over the duration of a culture.

The aim of this work was to exploit the fact that phase contrast microscopes are nowadays typically equipped with digital cameras, which enable the acquisition of high-resolution images. By automating the analysis of phase contrast microscopy (PCM) images, cost-effective non-invasive methods for the generation of quantitative cell culture data could be devised. The method should be compatible with existing cell culture ware, and it should be quick so that it does not introduce a bottleneck in existing cell culture protocols. It should also produce high quality results regardless of the imaging conditions, operator, microscope manufacturer, and cell line used. Such a method would thus enable inter- and intra-laboratory reproducibility. Finally, it should be accessible to researchers from a non-technical background, and not require extensive calibration or parameter tweaking for its use. In other words, the time to first result (TTFR) should be as short as possible.

The objective was therefore to design, implement and thoroughly validate a segmentation algorithm for adherent cells in PCM images that meets the requirements set out above. In the context of this chapter, segmentation refers to the classification of image pixels as either foreground (i.e. cells) or background. The aim was not the segmentation of individual cells as this task would be nearly impossible for colony-forming cell lines without the use of high contrast makers (e.g. fluorescent probes). For instance, mouse embryonic stem cells (mESCs), which were used as a model for the validation of the algorithms presented in this chapter, start forming colonies shortly after seeding. Within a colony, the boundaries between cells are very difficult to detect, preventing the segmentation of individual cells solely based on PCM images

(Figure 2-1). With the view of developing an algorithm as broadly applicable as possible, the focus was set on the segmentation of cellular objects, encompassing both individual cells (e.g. at the early stages of a cell culture) and colonies.

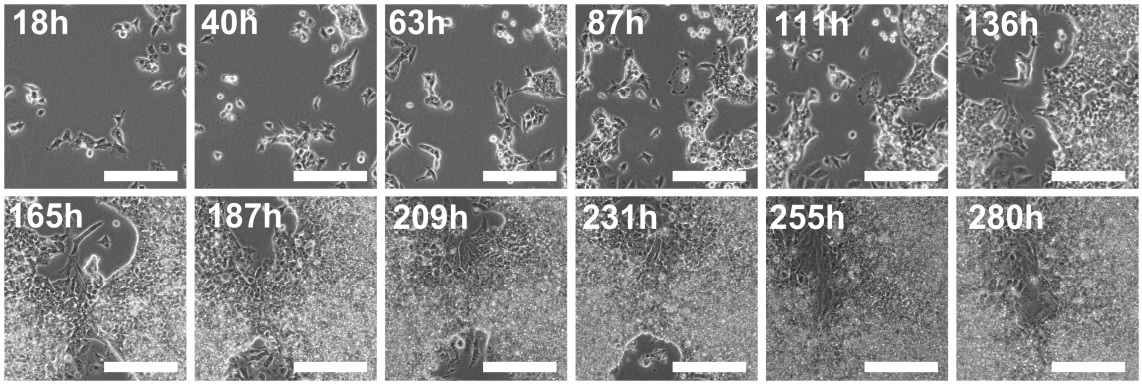


Figure 2-1 Visual features of mouse embryonic stem cells (mESCs) at various stages of a low seeding density expansion experiment. Individual cells can initially be easily distinguished, but rapidly form dense colonies where boundaries between cells are no longer visible. A low seeding density was used to show the whole range of cell visual features but colonies would appear earlier in a typical experiment with a higher seeding density. Acquired using a 10× objective. Scale bars are 200µm.

To best tackle cellular object segmentation in PCM images, multiple algorithms were considered. They all shared a common underlying concept: segmentation based on the degree of pixel intensity homogeneity. Indeed, cell regions usually present high variations in intensity (i.e. low homogeneity) due to the fine structures that make up cells (e.g. cell membrane, nucleus), as opposed to much more homogeneous background regions (see section 2.1.2 in this chapter, and 1.3.2 in chapter 1 for further details). In this thesis book, algorithms based on this principle are called contrast filters. Those were chosen as a starting point based on performance considerations, and more specifically their low processing times when compared to more complex approaches (see section 1.3.2).

This chapter focuses on the description and thorough evaluation of different segmentation algorithms based on contrast filters. Furthermore, the use of dynamic programming to further increase segmentation performance is investigated. Robustness to changes in imaging conditions as well as generalisation to multiple cell lines, microscope models and camera types are also assessed. Examples of applications to adherent cell culture monitoring will be presented in chapter 3.

2.1 Description of PCM segmentation approaches based on contrast filters

This section introduces the different steps making up the PCM image processing algorithms investigated in this chapter, including image pre-processing, segmentation, and *post-hoc* correction of halo artefacts. In particular, three types of contrast filters that facilitate the segmentation of PCM images are described and illustrated using simple examples. The performance of these approaches is thoroughly assessed in section 2.3.

Unless otherwise specified, MATLAB (version 2010a up, MathWorks, USA) containing all available toolboxes was used for image processing. In particular, this work made extensive use of the image processing toolbox.

2.1.1 Image pre-processing

PCM images were stored using uncompressed TIF or PNG formats. Depending on the microscope setup used (see section 2.2.2), images were either grayscale or RGB. The latter were first converted to grayscale using MATLAB's *rgb2gray* function, which computed a weighted average of the RGB channels (weighted 0.30, 0.59 and 0.11 for red, green and blue channels respectively). Images were then converted to a double representation using the *im2double* function. This operation rescaled the data so that pixel intensities ranged from 0 to 1.

2.1.2 Segmentation

2.1.2.1 Intensity thresholding

The segmentation of a digital image consists in classifying each pixel as either foreground (objects of interest) or background. In many cases, a simple intensity threshold can be sufficient to obtain good segmentation results. This is done by settings all pixels with an intensity above a given threshold (ε) to 1, otherwise to 0, resulting in a binary image G (Equation 2.1).

$$G(x,y) = \begin{cases} 1 & \text{if } C(x,y) > \varepsilon \\ 0 & \text{if } C(x,y) \leq \varepsilon \end{cases} \qquad 2.1$$

This operation was carried out in MATLAB either using the built-in *im2bw* function or by taking advantage of logical indexing.

PCM images are rarely bi-modal, in other words there are not two well-separated peaks representing the cell and background pixels (Figure 2-2.A). Instead, there is a rather large overlap in intensity between background and cell pixels (Figure 2-2.B).

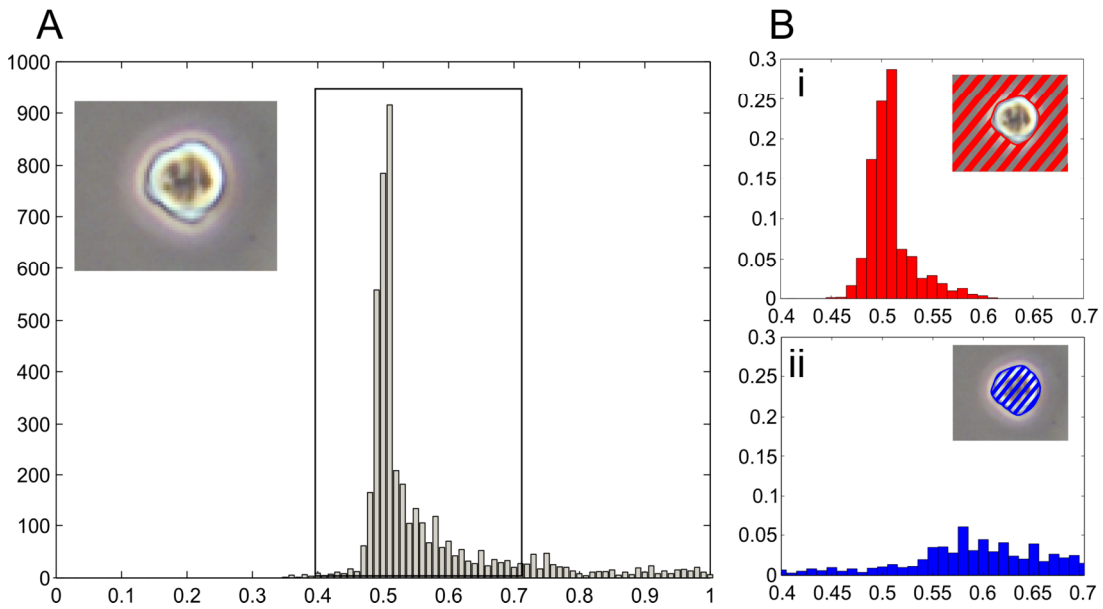


Figure 2-2 Intensity histogram of a PCM image of a mES cell shortly after seeding. (A) Histogram of the whole image, with an insert of the corresponding PCM image. The rectangle denotes the region spanned by the zoomed-in normalised histograms shown in (B). (B) Normalised intensity histograms for the background (i) and the cell pixels (ii).

Consequently, simple intensity thresholding usually produces unsatisfactory results (Figure 2-3). Indeed, there is either a significant number of background pixels detected as cells (i.e. large number of false positives) or a significant part of the cell missing (large number of false negative pixels).

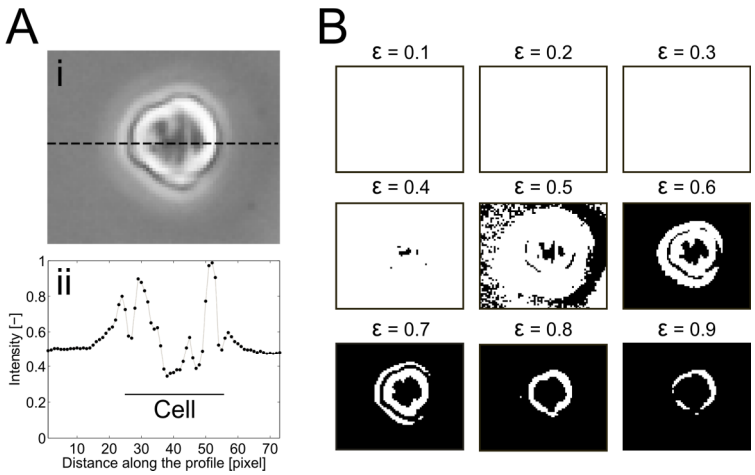


Figure 2-3 Simple intensity threshold for PCM image segmentation. (A) Typical intensity profile of a PCM image of a single mES cell shortly after seeding. (i) Grayscale PCM image overlaid with a dashed line indicating the pixels used to build the intensity profile. (ii) Intensity profile along the dashed line in (i). Insert text denotes the section of the profile corresponding to the cell. (B) Binary images obtained after thresholding at the indicated threshold values (ε).

It is thus necessary to pre-process the image to facilitate its subsequent thresholding. One approach consists in taking advantage of the difference in the homogeneity of pixel intensities between the image background and cellular objects (Bradhurst et al., 2008; Juneau et al., 2013; Topman et al., 2011). The range of intensities for background pixels is usually relatively narrow (Figure 2-2.B.i) when compared to the very broad distribution of cell pixels (Figure 2-2.B.ii).

Local variation in intensity can be detected using various filters, such as local standard deviation and range filters.

2.1.2.2 Hard-edged local standard deviation filter (heSTDEV)

A local standard deviation image *S(x,y)*, whose pixel values correspond to the standard deviation in *M*×*N* windows centred at each pixel locations of the image to filter *I*, can be computed efficiently using a convolution scheme according to the formulation of standard deviation based on the expected value of a random variable *E[X]* (equation 2.2). Alternative approaches for the computation of the standard deviation can be found in annex E.

$$\sigma = \sqrt{E\left[X^2\right] - \left(E\left[X\right]\right)^2} \tag{2.2}$$

When using a kernel whose all elements are equal to $\frac{1}{M \times N}$, convolution corresponds to the computation of local intensity means, which are equivalent to expected values. The standard deviation image can thus be computed as follows:

$$S(x,y) = \sqrt{\frac{H \otimes I^2}{n-1} - \frac{\left(H \otimes I\right)^2}{n\left(n-1\right)}} \tag{2.3}$$

where $\otimes$ denotes the convolution of the image *I* by the kernel *H*, which is a matrix of dimensions M×N whose *n* elements are equal to 1. For this work, the *stdfilt* method in MATLAB was employed.

Examples of intensity profiles obtained after filtering using a hard-edged local standard deviation filter are shown in Figure 2-4.A. In addition, binary images obtained by thresholding the filter output are shown in Figure 2-4.B.

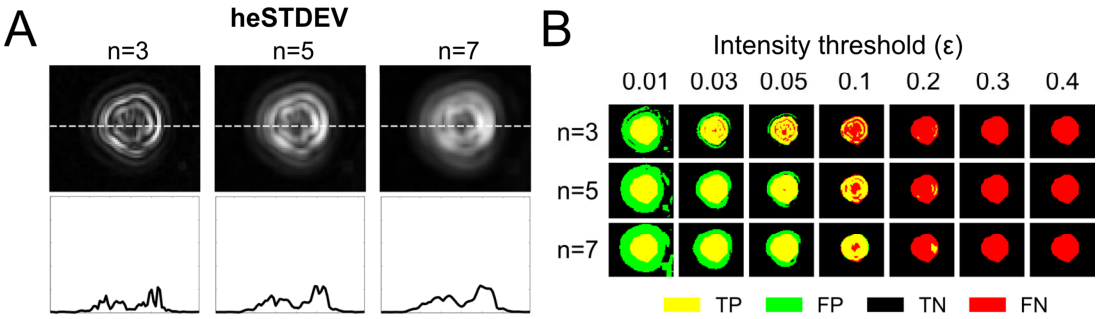


Figure 2-4 Hard-edged local standard deviation (heSTDEV) filter output for a PCM image of a mES cell shortly after seeding. (A) Intensity profile. The top image shows the filter output, the bottom image is the intensity profile along the dashed line on the filter output. (B) Threshold applied to the heSTDEV filter outputs. The resulting binary image is compared to a manually delineated ground truth image. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN). TP, FP, TN, and FN are defined in section 2.2.4.

2.1.2.3 Normalised soft-edged local standard deviation filter (nseSTDEV)

Extensive preliminary work was done in order to find viable alternatives to hard-edged local contrast filters. The best candidate was a normalised soft-edged local standard deviation where the hard-edged sliding-window was replaced by a soft-edged window computed using a Gaussian kernel. The filtered image *C(x,y)* can be computed as following:

$$C(x,y) = \frac{\sqrt{G_\sigma \otimes I^2 - \left(G_\sigma \otimes I\right)^2}}{G_\sigma \otimes I} \qquad 2.4$$

where $\otimes$ denotes the convolution of the image I with a Gaussian kernel G. Gaussian functions are radially symmetric and separable (Sahoo, 1990). In other words, a relatively complex 2D convolution can be carried out in two steps, by computing the convolution of the image with a 1-dimensional kernel in both horizontal and vertical directions. The kernels are built by sampling the non-normalised Gaussian function of standard deviation σ:

$$G_\sigma(x) = e^{\frac{-x^2}{2\sigma^2}} \qquad 2.5$$

In order to avoid excessive darkening or brightening of the image, the normalised kernel K_{Norm} is computed as follows:

$$K_{Norm} = K./\sum_{k=1}^{M}\sum_{m=1}^{N} K\left(x_k, y_m\right) \qquad 2.6$$

where ./ denotes an element-wise division.

The size of the window used for the local contrast filter depended on the standard deviation of the Gaussian kernel G. In order to avoid cutting off of significant values, the dimension of the 1-dimensional kernel was $2 \times 3\sigma + 1$, in other words 3σ on each side of the central point.

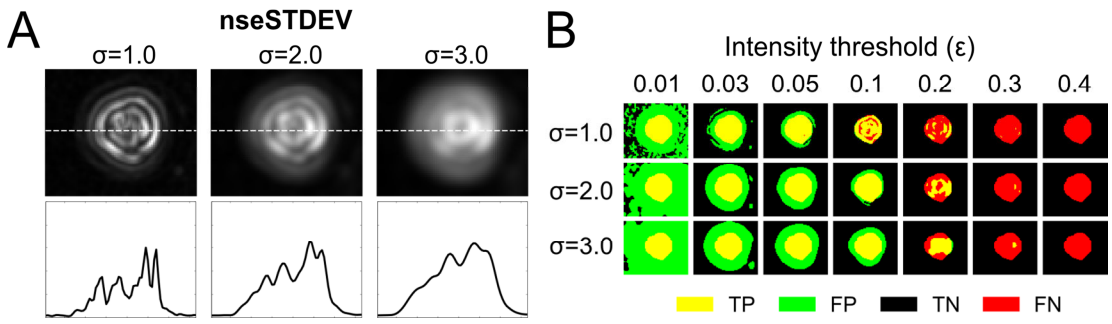


Figure 2-5 Normalised soft-edged local standard deviation (nseSTDEV) filter output for a PCM image of a mES cell shortly after seeding. (A) Intensity profile. The top image shows the filter output, the bottom image is the intensity profile along the dashed line on the filter output. (B) Intensity threshold applied to the nseSTDEV filter outputs. The resulting binary image is compared to a manually delineated ground truth image. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN). TP, FP, TN, and FN are defined in section 2.2.2.1.

Examples of intensity profiles obtained after filtering using a normalised soft-edged local standard deviation filter (nseSTDEV) are shown in Figure 2-5.A. In addition, binary images obtained by thresholding the filter output are shown in Figure 2-5.B.

2.1.2.4 Range filter (RANGE)

Range filters set the value of each pixel of an image to the difference of the maximum and minimum intensity values within its neighbourhood (Juneau et al., 2013). This can be efficiently determined using grey-scale morphology operations by computing the difference between the eroded (minimum within the neighbourhood) and dilated (maximum within the neighbourhood) images (Gonzalez and Woods, 2007). In MATLAB, this was done by calling the *rangefilt* function.

Examples of intensity profiles obtained after filtering using a range filter are shown in Figure 2-6.A. In addition, binary images obtained by thresholding the filter output are shown in Figure 2-6.B.

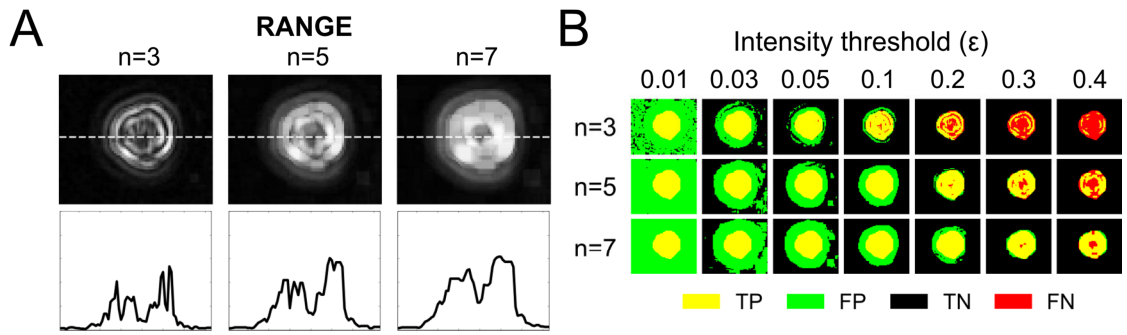


Figure 2-6 Range filter output for a PCM image of a mES cell shortly after seeding. (A) Intensity profile. The top image shows the filter output, the bottom image is the intensity profile along the dashed line on the filter output. (B) Intensity threshold applied to the RANGE filter outputs. The resulting binary image is compared to a manually delineated ground truth image. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN). TP, FP, TN, and FN are defined in section 2.2.4.

2.1.3 Post-processing hole filling and small objects removal

Binary morphological operations were used to enhance the binary mask obtained after contrast filtering and thresholding. Objects present in the binary mask were located using connected-components analysis (i.e. detection of all distinct groups of pixels that are connected). MATLAB's *regionprops* and *bwconncomp* functions were used for this purpose. Small object removal was achieved by removing (i.e. set to a value of 0) all pixels associated with objects of area less than a threshold parameter A_{RSO}. Similarly, holes were detected by locating connected-components on the complement of the binary image (i.e. background becomes foreground and vice versa). Holes smaller than a threshold A_{FOH} were filled. Unlike some approaches that were previously described, it was not possible to use the simpler flood fill method (i.e. fill any region that does not touch the border of the image) as cellular objects in PCM images can have sizeable holes that should be detected as background rather than cell pixels.

2.1.4 *Post-hoc* halo correction

It was previously suggested that the transition from halo artefact to cell membrane pixels is accompanied by changes in the direction of the intensity gradient (Bradhurst et al., 2008). A *post-hoc* halo artefact correction method based on this observation was developed. The main assumption was that segmentation by thresholding of the contrast filter output would almost always correctly detect cell pixels (high recall), but with portions of the halo artefacts incorrectly classified as cell pixels (decreased precision).

Halo artefacts usually bleed into the background of the image, resulting in a steady increase in pixel intensity when transitioning from the unaffected background to the halo region. This intensity is usually maximal at the cell border. The gradient vector field thus tends to point towards the cell, until its direction suddenly reverses when reaching the dark cell membrane (Figure 2-7.A). Thresholding of the contrast filter output usually produces a mask whose outer pixels fall within the region between the start of the halo artefact and the cell membrane. The

method proposed here thus makes use of that assumption, to correct the binary mask, and to make it conform to the cell contour.

The first steps involved the computation of the intensity gradient of the PCM image. The various approaches that can be used for this computation are presented and discussed in annex B. After preliminary tests, it was decided to use the compass operators approach (see annex B.1.2). In brief, the image was convolved using 8 kernels that were sensitive to directions in increments of 45°. The magnitude of the gradient was given by the largest response obtained after the 8 convolutions and the direction by the kernel that produced said response. The set of kernels used were the Kirsch compass kernels (see Annex B.1.2). Unlike other approaches that usually require the determination of the gradient direction from the magnitude vector, the compass operators method yielded the direction in a form that was directly usable by the halo correction algorithm (i.e. quantized for the 8 cardinal directions).

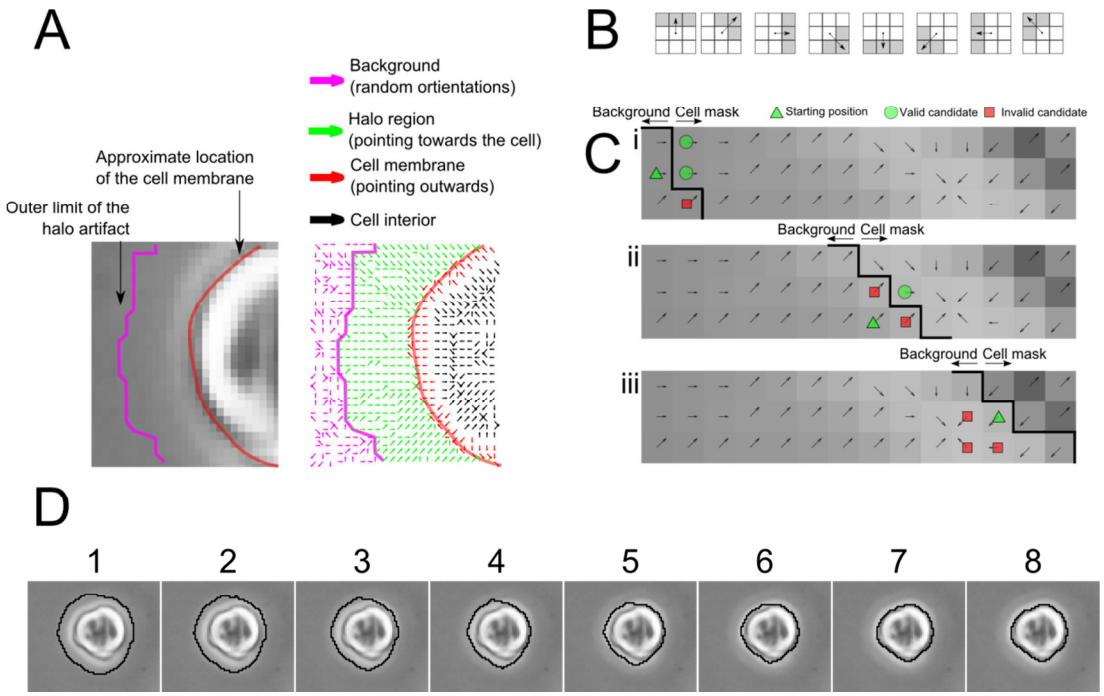


Figure 2-7 Method for *post-hoc* correction of halo artefacts. (A) Typical image gradient in the neighbourhood of a cell. (B) In addition to the gradient direction, the algorithm also considers the two nearest directions. (C) Rules to determine whether a pixel should be removed from the binary mask or not. (D) Raw PCM image overlaid with the border of the binary mask as it is shrunk over 8 iterations.

The iterative algorithm itself was first prototyped in MATLAB and then re-implemented in C++ for faster computation. A detailed description of the algorithm is shown in **Listing 2-1**.

In short, the algorithm is first initialised with a list of pixels to visit, which corresponded to the outer border of the binary mask obtained after thresholding of the contrast filters output. For each of the pixels in that list, the algorithm checked whether the gradient pointed inside the mask or outwards. If it was the latter, nothing was done as the mask was only allowed to shrink, not expand. If the neighbouring pixel that the gradient pointed to was in the binary mask, said neighbouring pixel was added to the list of locations to visit during the next iteration and the

current pixel on the mask was set to 0. The same process was repeated for the two closest directions to that of the gradient (e.g. if the gradient pointed north-west, both north and west directions were also considered, see Figure 2-7.B). Essentially, the binary mask was iteratively shrunk as long as the gradient pointed inwards. The process stopped when it reached the point where the direction of the gradient was inverted (i.e. pointing outwards), which corresponded to the cell membrane (Figure 2-7.C and D).

Listing 2-1 Details of the iterative halo correction algorithm. *I* is the input PCM image, *B* is the binary image resulting from the intensity thresholding of the contrast filter output.

```
01. imgGradient = compute intensity gradient of input image I
02. Initialise pixelsToVisit with the coordinates of the object
    boundaries pixels in B
03. While pixelsToVisit is not empty
04.     For each location in pixelsToVisit
05.         Lookup the gradient direction at location from imgGradient
            as well as that for the two closest directions

06.         For each pixel pointed at by the three directions
07.             If pixel in binary mask B
08.                 Add pixel to pixelsToVisitNext
09.                 Add location to maskPixelsToDelete
10.     Set all pixels listed in maskPixelsToDelete to 0 in B
11.     Set pixelsToVisit to pixelsToVisitNext
```

It was necessary to implement safeguards to prevent the halo correction algorithm to have a detrimental effect in a few edge cases. A parameter R_{max} was used to control the fraction of an object that could be shrunk during the iterative process. For instance, if R_{max} was set to 0.3, only 30% of the total area of an object could be shrunk during correction. Such a value would result in effective halo artefact correction in most cases but would prevent objects to be incorrectly removed from the binary mask if the intensity profile in their neighbourhood did not conform to what was expected. Typical situations included ones where the halo was either non-existent or discontinuous, possibly resulting in the intensity gradient pointing inwards without ever changing direction. The incidence of such cases was found to be very low but this safeguard was implemented to improve the generalisation of the approach.

2.2 Methods for dataset generation and segmentation performance evaluation

2.2.1 Cell culture

2.2.1.1 Mouse embryonic stem cell

Mouse embryonic stem cells (E14Tg2a, Oct4-GiP, passage number < 70) were kindly donated by Stem Cell Sciences (Cambridge, UK). For maintenance, cells were grown in T-25 flasks

(Nunc, USA) pre-coated with 0.1% gelatin. The working volume of culture medium (see Table 2-1 for formulation) was 5 ml. The culture medium was exchanged daily.

Table 2-1 Embryonic stem cell maintenance medium formulation

Component	Concentration
DMEM KnockOut (Invitrogen, UK)	-
β-mercaptoethanol (VWR, UK)	0.1 mM
Fetal Bovine Serum (FBS, Invitrogen, UK)	15% v/v
Non-essential amino acids (Invitrogen, UK)	1% v/v
sL-Glutamine (Invitrogen, UK)	2 mM
Sodium pyruvate (Invitrogen, UK)	1 mM
Leukemia inhibitory factor (LIF, Millipore, UK)	10^3 Units ml^{-1}

Flasks were kept in an O_2/CO_2 incubator (Sanyo Biomedical Division, UK) at 37°C and 5% CO_2. Cells were passaged every 3 days at a 1:10 (area/area) split ratio: supernatant was discarded and cells were washed with 2 ml of Dulbecco's Phosphate Buffered Saline (DPBS) without Ca^{2+}/Mg^{2+} (Invitrogen, UK). Cells were then incubated for 3 minutes with 500 µl of a 0.25% (v/v) trypsin / 0.53 mM EDTA solution (Invitrogen, UK). After quenching with 2 ml of fresh culture medium, cells were centrifuged for 3 minutes at 300 g, re-suspended in fresh culture medium and transferred to a new gelatin-coated flask.

2.2.1.2 Chinese hamster ovary cells and Human neuroblastoma cell

Chinese hamster ovary cells (CHO-K1, ATCC CCL-61) and human neuroblastoma cells (SK-N-SH, ATCC CRL-2266) were cultured and maintained in T-25 flasks (Nunc, UK) following the protocol outlined in 2.2.1.1, with the only deviation being the formulation of the culture medium which is shown in Table 2-2.

Table 2-2 CHO and Neuroblastoma cell culture medium formulation

Component	Concentration
DMEM High-Glucose (Invitrogen, UK)	-
β-mercaptoethanol (VWR, UK)	0.1 mM
Fetal Bovine Serum (FBS, Invitrogen, UK)	10% v/v
Non-essential amino acids (Invitrogen, UK)	1% v/v
L-Glutamine (Invitrogen, UK)	2 mM

2.2.2 Image acquisition

Unless specified otherwise, the cells used for image acquisition were cultured in 6-well plates (Nunc, UK). Multiple imaging setups were used during the course of this project (Table 2-3). Unless otherwise specified, the results presented were acquired with setup no. 2, using a 10× objective and phase contrast illumination. Each image then corresponded to a field of view of 1.27 mm x 0.95 mm, i.e. approximately an area of 1.20 mm^2.

Table 2-3 Various imaging setups used during the course of this project				
Setup	Microscope	Camera	Image	Acquisition Software
1	Nikon TE-2000	Basler Scout	Grayscale 1280x960	Vision Assistant (National Instrument)
2	Nikon Ti-E	Nikon DS-Fi1	Colour 1280x960	NIS Elements-F (Nikon)
3	Olympus IX71	Hamamatsu ORCA-ER	Grayscale 1344x1024	AxioVision (Olympus)
4	Zeiss Axiovert 135	Hamamatsu ORCA-R2	Grayscale 1344x1024	Volocity (PerkinElmer)

2.2.3 Datasets for performance evaluation

PCM images were manually annotated using the Paint.NET image-editing tool (dotPDN LLC, USA). Upon loading of the raw PCM image, two additional layers were added. One was used to delineate the cells while the other would represent the background. The resulting binary images were then saved in an uncompressed format (either TIF or PNG).

The main dataset used for performance evaluation consisted of 50 mouse embryonic stem cell (mESC) PCM images. Those images were acquired during the course of multiple expansion experiments in imaging conditions representative of a typical cell culture laboratory (image setup 2 in Table 2-3). The cells were cultured in 6-well plates. The set of images spanned the whole range of confluencies (i.e. cell coverage), thus capturing the full range of appearances typically observed during the course of a culture (Figure 2-8.A). The image size was limited to 250×250 pixels to facilitate their manual annotation and minimise the computational complexity of the parameter space screening process (Figure 2-8.B).

A dataset for the evaluation of generalisation to other cell lines was also generated. It consisted of six 500×500 PCM images, three of Chinese hamster ovary cells and three of Human neuroblastoma cells. In addition, PCM image of other cell types were obtained from public databases (e.g. Cell: An Image Library[2]) or other researchers.

[2] http://www.cellimagelibrary.org/ - Last accessed 20.06.2014

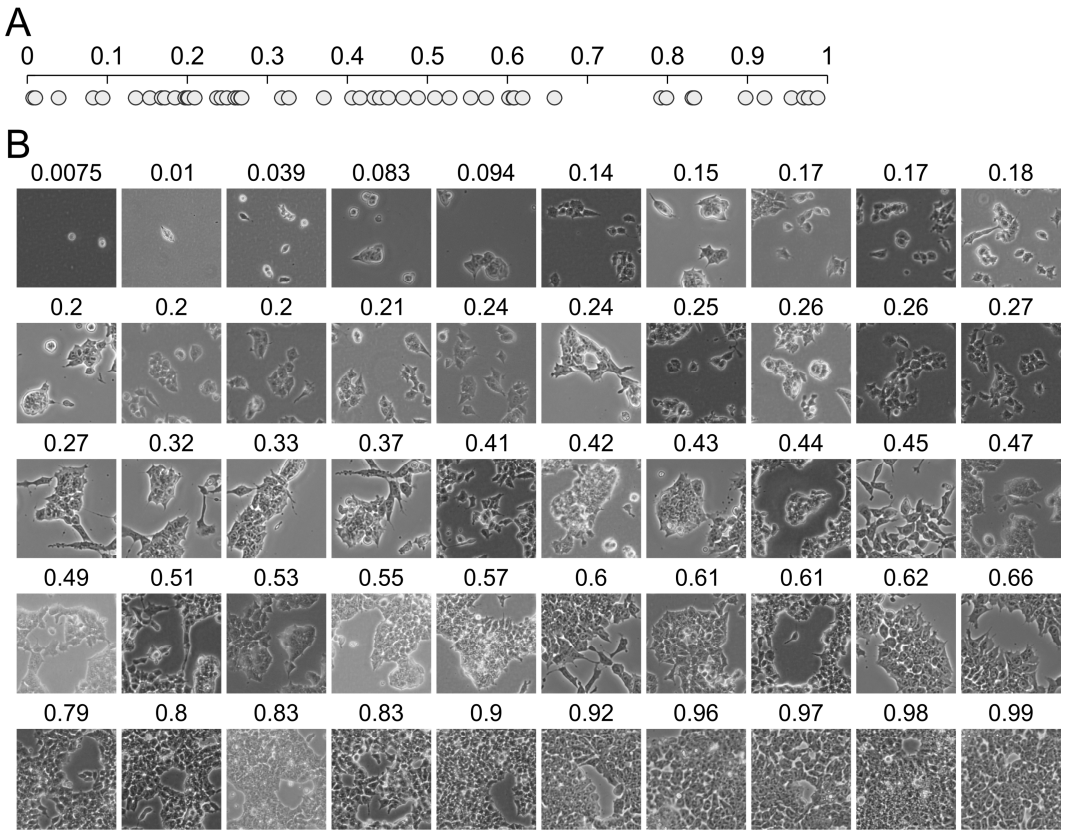


Figure 2-8 Mouse embryonic stem cell (mESC) PCM image dataset. (A) Range of confluencies covered by the set (as computed from the manually annotated images). (B) Images of the set ordered by confluency (increasing left to right, top to bottom). The text above each image denotes the confluency of the corresponding image.

Finally, a dataset was generated for the evaluation of the sensitivity of segmentation algorithms to changes in imaging conditions. Three fields of view of a mESC culture in a 6-well plate were imaged using four combinations of microscopes, cameras and acquisition software (Table 2-3).

2.2.4 Evaluation of segmentation performance

Segmentation performance was primarily determined by comparison of the segmentation outcomes to ground truth images that were manually annotated by a human expert. A pixel-wise comparison of the segmentation output with the ground truth images, both being binary images, yielded a so-called confusion matrix (Table 2-4).

Table 2-4 Confusion matrix for the comparison of segmentation outputs with ground truth images

		Segmentation output	
		Cell	**Background**
Ground truth image	**Cell**	True positive (TP)	False negative (FN)
	Background	False positive (FP)	True negative (TN)

Various metrics can be used to summarise the confusion matrix (Fawcett, 2006). They are often referred to as receiver operator characteristics (ROCs), and are described in Table 2-5.

Table 2-5 Receiver operator characteristics (ROCs) for the analysis of the confusion matrix

Metric	Expression	Interpretation
Accuracy	$\frac{TP + TN}{TP + FN + FP + TN}$	Fraction of pixels correctly labelled by the segmentation algorithm
Precision	$\frac{TP}{TP + FP}$	Fraction of pixels labelled as cells that are actually cell pixels
Recall	$\frac{TP}{TP + FN}$	Fraction of cell pixels correctly labelled
F-score	$\frac{2 \times TP}{FP + FN + 2 \times TP}$	Also known as the F-measure, F-value, and Dice coefficient. Measure of agreement (overlap) between the output and the ground truth image. Values of 0 and 1 signify no and complete overlap respectively

To illustrate the relevance of the various ROCs, a simple artificial example was analysed. Binary images of concentric circles simulated the manual annotation of an expert and the segmentation output of an algorithm (Figure 2-9.A). The circle representing the ground truth data had a radius 50% larger than that of the simulated segmentation output. The smaller circle was progressively shifted horizontally pixel by pixel, effectively worsening the performance of the simulated segmentation algorithm (Figure 2-9.B).

Even though the simulated circle was not large enough to cover the whole ground truth circle, the precision was maximal as long as the smaller circle was entirely contained within the larger one, as false negatives are not taken into account for the calculation of the precision. The recall value was initially 44%, or approximately the ratio of the area of the small circle to that of the large circle. It essentially showed that 44% of the cell pixels were correctly detected. The F-Score, as a measurement of the overlap between the simulated output and the ground truth data, remained constant at ~60% as long as the simulated output was contained within the ground truth circle, before dropping and reaching zero as both circles were completely separated. In contrast, accuracy started at 90% and only slowly decreased, reaching about 80%

when the two circles were completely separated. It remained high as a large number of background pixels were still correctly labelled, even though the object of interest was completely missed.

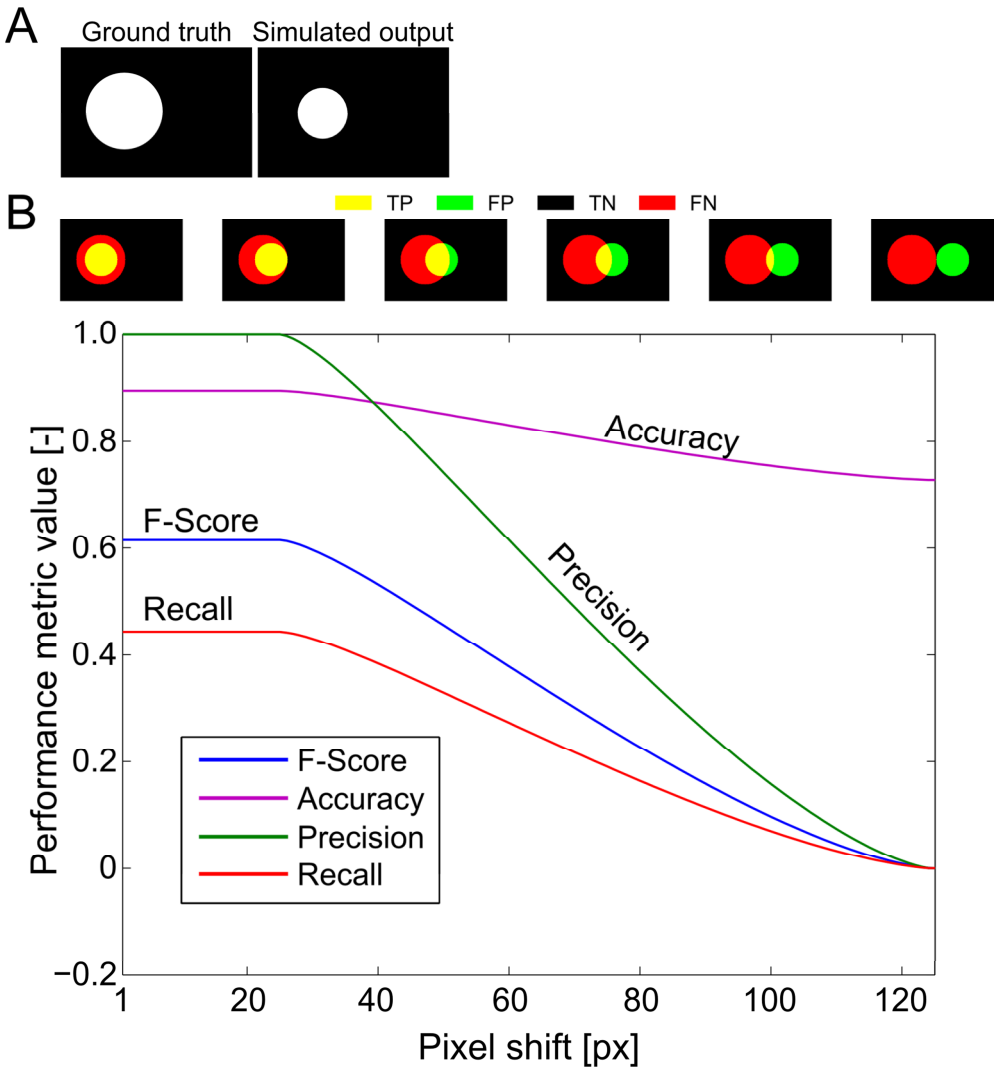


Figure 2-9 Synthetic example illustrating the behaviour of the various ROCs considered. (A) Segmentation output and ground truth data are both simulated by circles of differing radii. (B) The circle representing the segmentation output is shifted to the right pixel by pixel right to simulate the worsening of segmentation quality.

Accuracy was thus inadequate for segmentation performance assessment in most cases. Accuracy is only a pertinent metric if the classes are perfectly balanced, i.e. if the number of cell pixels is the same as the number of background pixels, which is rarely the case for microscopy images. F-score was found to be much more suitable for segmentation evaluation. Similar results were obtained in the case where the simulated algorithm output was smaller than the simulated ground truth. More specifically, both the F-score and accuracy behaved identically to the case presented above (i.e. simulated output smaller than simulated ground truth) while the profiles for recall and precision were switched around.

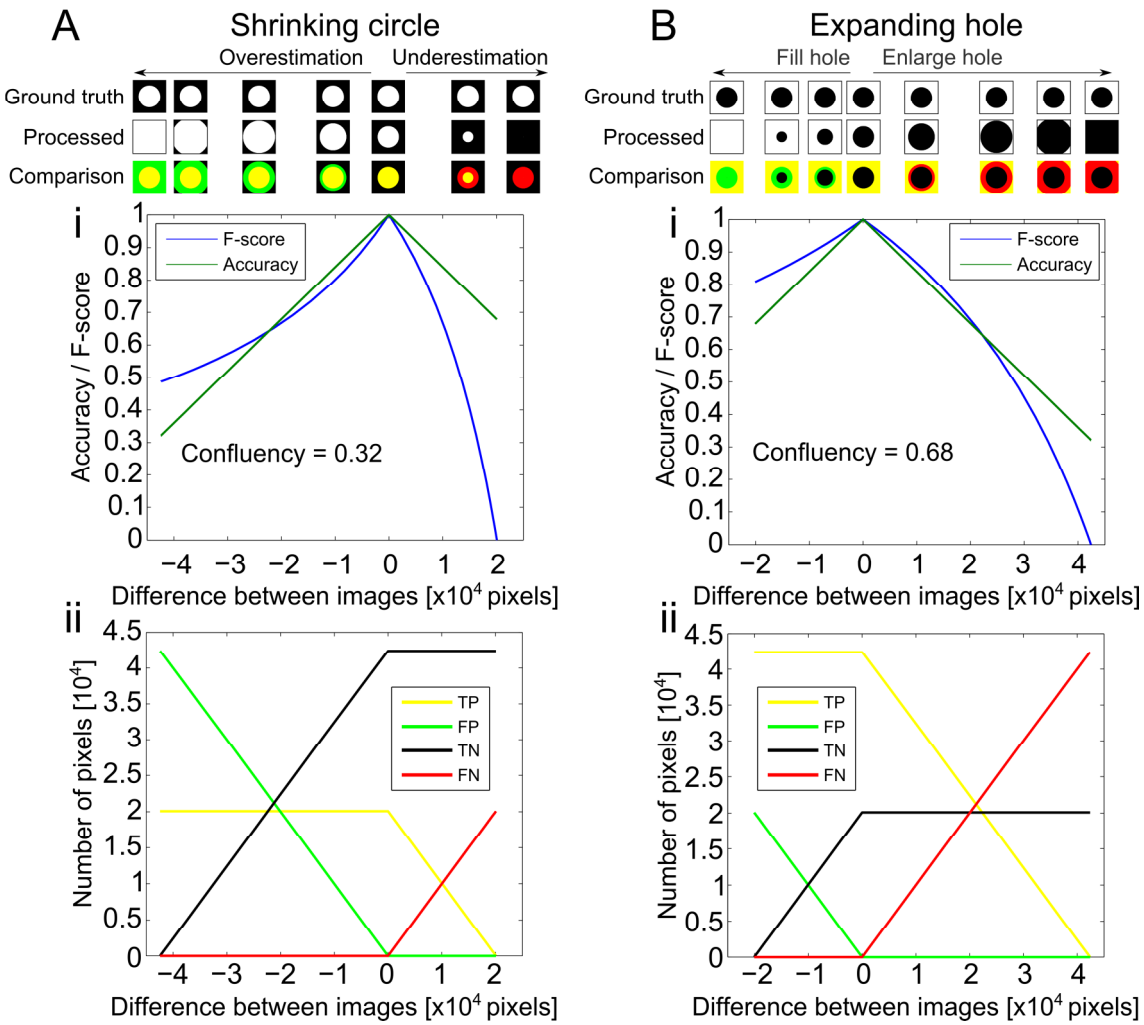


Figure 2-10 Detailed analysis of the F-score behaviour using artificial examples. (A) Artificial example to determine the effect of overestimation or underestimation of the area of the object of interest (e.g. a single cell). Ground truth data was represented by a circle of constant radius. The simulated segmentation output was abstracted by a second circle whose radius was varied so that it would cover the whole image, be identical to the ground truth circle or shrink until it disappeared. In all cases, the x-axis shows the difference in cell pixels (pixels with a value of 1) between the artificial ground truth and processed images. (i) Effect on F-score and Accuracy. (ii) Effect on true/false positives and true/false negatives (B) Artificial example to determine the behaviour of the F-score and accuracy when assessing the segmentation performance of an object with a hole. The ground truth was represented by a filled square with a circle hole. The size of the hole was varied, from no hole to one that covered the whole image. (i) Effect on F-score and Accuracy. (ii) Effect on true/false positives and true/false negatives.

As a metric, the F-score also has some limitations and counter-intuitive behaviours that are important to keep in mind when it is used to assess segmentation performance. Simple artificial examples illustrated some of these limitations in scenarios that were relevant to microscopy image processing. In the situation where a single cellular object is imaged, the changes in F-score associated with overestimation or underestimation of the cell area were not identical (Figure 2-10.A.i). When overestimating the cell area, the F-score decreased slowly as the true negatives went down, the false positives went up and the true positives remained constant, essentially preserving the overlap between segmentation output and ground truth (Figure 2-10.A.ii). In contrast, an underestimation of the object area resulted in a sharp and rapid decrease in F-score. This was mostly caused by the decrease in true positives, as less and less cell pixels were correctly detected. Another artificial scenario investigated the impact of varying segmentation performance in the case of an object that contained a sizable hole. Again F-score

was only mildly impacted when the size of the hole was underestimated as it preserved the number of true positives whereas overestimating the size of the hole resulted in a much sharper drop due to a decreasing number of true positives and a lower degree of overlap (Figure 2-10.B.i and Figure 2-10.B.ii).

These two scenarios were further studied for various degrees of image confluency, i.e. the number of positive pixels in the ground truth (Figure 2-11). For both cases, the impact on the F-score error was much more pronounced for lower confluency values. In other words, the impact of segmentation error on the F-score was found to be biased by the number of cell pixels in an image. The F-score for low confluency microscopy images (e.g. shortly after seeding) will therefore be more sensitive to segmentation errors than that for images acquired at later time points (intermediate to high confluency).

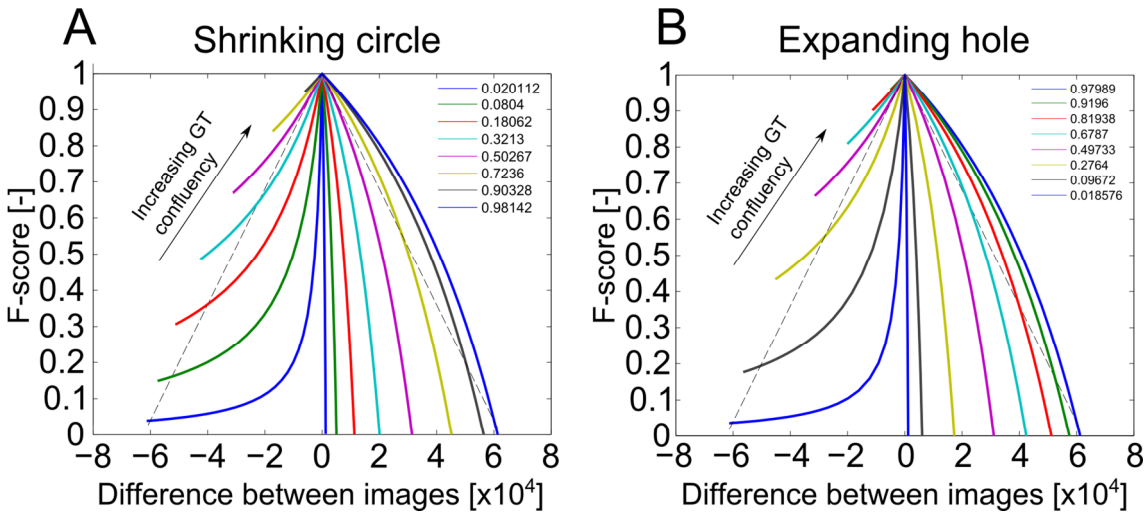


Figure 2-11 Further investigation of the F-score behaviour. Scenarios are those presented in Figure 2-10 for various degrees of confluency (i.e. the size of the ground truth object (A) and hole (B) was varied, while maintaining the same image size).

2.2.5 Optimisation of segmentation parameters

Determining optimal segmentation parameters is often as important as devising the underlying image processing algorithm. Parameter tuning can be challenging as the chosen values should generalise to unseen data and consistently maintain high processing performance regardless of the data set considered. Essentially, the goal of parameter tuning is to find values that will enable the mapping of an input image to a manually annotated ground truth with the aim to achieve the highest possible ROC values with the devised image processing algorithm (Figure 2-12.A). The process has to be systematic and cannot rely on manual tweaking. Because this project involved continually evolving algorithms that would require evaluation and optimisation very often, a specialised framework was built for this task. The objective was to develop a tool that would enable the rapid screening of parameter spaces with potentially millions of possible combinations.

2.2.5.1 *Screening-based parameter optimisation*

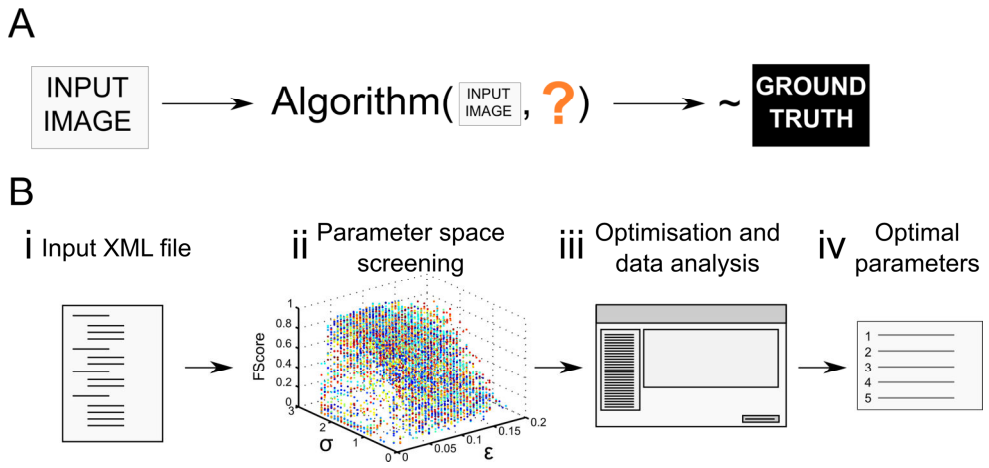


Figure 2-12 Parameter tuning of the segmentation algorithm. (A) General idea behind parameter tuning: find the parameters (here denoted by ?) so that the given algorithm can be used to map the input image to the ground truth image as closely as possible. (B) Schematic showing a typical optimisation procedure. (i) A user-generated XML input file is loaded and parsed by the framework. (ii) The parameter space is screened efficiently using parallel computations. In this example, the contrast filter operation was optimised by varying the σ and ε parameters. (iii) The user interface is used to explore the data and initialise optimisation using various built-in methods. (iv) The optimal parameters are determined.

Considerations made for the parameter tuning tool and their implementation can be found in annex C. In brief, the framework was built using MATLAB's object-oriented capabilities combined with a scripting language (based on Extensible Markup Language, or XML) and a graphical user interface (GUI). First, an XML file was created by the user in which the various elements of the image-processing algorithm to optimise were broken down into individual operations (Figure 2-12.B.i). These operations could output an image, numerical data or a combination of both. The input parameters (which were images, numerical values or string of characters) were defined either as a single value, as multiple values or as a range of values (for numerical parameters only). These options were parsed by the framework and were used to generate the parameter space to investigate, which could contain millions of combinations (Figure 2-12.B.ii). The framework screened each combination of parameter values through a grid search and scored them for each input image using the metrics described in section 2.2.4. As the framework was designed with parallelisation in mind, increasing the number of cores available drastically reduced the time required to screen the parameter space. The GUI could then be used to determine the optimal segmentation parameters based on the data generated during the screening process (Figure 2-12.B.iii). There were three built-in optimisation functions to determine the optimal parameter values for a set of input images:

- ***Per-image optimisation***: the optimisation was done independently for each image. It essentially selected the parameter set that produced the lowest segmentation error (as determined using the specified optimisation metric) for any given image. This informed on the best overall performance that can be achieved for the set of input images using the chosen operations.

- ***Whole-set optimisation***: the entire image set was used for the determination of the optimal parameter values. The segmentation errors were summed across all input images and the parameter set associated with the lowest value was chosen as the best overall parameter set. This optimisation method gave a sense of what performance could be achieved for the set of input images using a single, common set of parameter set values. As a result, the performance for certain images could potentially be compromised in order to obtain better overall results.

- ***Leave-one-out cross-validation***: this was an iterative optimisation process. During a given iteration, all images but one (i.e. the left out image) were used to determine optimal parameter as described for the whole-set optimisation approach. The segmentation performance was then evaluated using the left-out image. This process was repeated until all images had been left out once. Leave-one-out cross-validation (LOOCV) was used to assess the generalisation of the optimal parameters to unseen images.

All three optimisation functions relied on the determination of the segmentation error. This was done using the metrics described in section 2.2.4. Instead of the absolute segmentation error, the framework employed an image-normalised segmentation error computed as following:

$$\overline{e_{i,j}} = \frac{e_{i,j}}{\max\left\{e_{i,k}, k = 1,2..,n\right\}} \tag{2.7}$$

where $\overline{e_{i,j}}$ was the image-normalised segmentation error for image i and parameter set j, $e_{i,j}$ was the absolute segmentation error and k varied from 1 to n denoting all the parameter sets screened. This was done in order to minimize image-specific bias when summing the segmentation error across all input images. By doing so, the parameter sets were ranked according to their relative performance for each image, which ensured that images that would always produce lower scores due to the complexity of the segmentation would contribute to the optimisation as much as the rest of the training set. Unless otherwise specified, the error was computed as:

$$e = 1 - Fscore \tag{2.8}$$

While the core code of the framework was kept the same for the duration of the project, it was continually extended by adding more and more operations (e.g. new contrast filters). This versatility enabled rapid validation of new ideas and concepts.

2.2.5.2 Optimisation using pattern search

This approach was conceptually different from screening-based optimisation. Instead of screening the entire parameter space, a pattern search algorithm was used to minimise an

objective function. Images were first labelled using an interactive painting tool. The labels could be sparse, in the sense that it was not necessary to annotate the whole image. When the optimisation was started, the Nelder–Mead method (also known as the downhill simplex method) was used to minimise the error function calculated as the mean of the error (computed as shown in equation 2.8) for all images considered for the optimisation. The segmentation error was only evaluated at the labelled locations instead of the full image. This approach was implemented in MATLAB using the *fminsearch* function.

2.3 Results of segmentation performance evaluation

The performance of various PCM image segmentation approaches (described in section 2.1) was first evaluated using the mESC dataset (Figure 2-8) before assessing their generalisation to other cell lines and imaging conditions.

2.3.1 Segmentation using contrast filters

The ideal algorithm produces high quality segmentations (e.g. as compared to human experts) in a time frame that makes its use in cell culture workflows practical. For this reason, the development of the segmentation algorithms focused on simple, computationally inexpensive filters and operations. The basic structure of the algorithms was based on a common image processing pipeline for microscopy image processing (Juneau et al., 2013): a contrast filter was first applied to highlight regions of high intensity variations that are, in the case of microscopy images, usually associated with the presence of cellular objects. The output of the filter was then thresholded to produce a binary mask, which was further refined using post-processing operations, including holes filling and the removal of small objects.

The contrast filters considered were a hard-edged local standard deviation filter (heSTDEV), a normalised soft-edged local standard deviation filter (nseSTDEV), and a range filter (RANGE). While the contrast filter step was different for all three algorithms, the rest of the operations (thresholding, hole-filling and small object removal) remained the same across the three approaches. For comparison purposes, an algorithm without a contrast filtering step was also considered by applying the intensity threshold directly to the raw PCM image.

Optimisation of parameter values was carried out using a grid search method. For each parameter, a range of values to evaluate was defined (Table 2-6). For all approaches compared, the parameter values were identical for the thresholding and post-processing operations. The parameters of the contrast filters were slightly altered due to intrinsic differences between the algorithms (e.g. window size can only be odd for RANGE and heSTDEV).

Table 2-6 Explored parameter space for segmentation algorithms evaluation (without halo correction).*

		Parameters		
Contrast filter	Operation	Symbol	Name	Screened values
None (Intensity only)	TR	A) ε	Binarisation threshold	0.01 to 1 (step = 0.01)
	PP	B) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		C) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
RANGE	CF	A) W	Window size	1 to 35 (step = 2)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
nseSTDEV	CT	A) σ	Gaussian kernel standard deviation	0.1 to 2.5 (step = 0.1)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
heSTDEV	CT	A) W	Window size	1 to 31(step = 2)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}

*LC, Local Contrast; STDV, local standard deviation; CF, Contrast Filter; TR, thresholding; PP, Post Processing

All combinations of parameter values within the parameter space considered were evaluated by comparison of the algorithm output with the corresponding manually annotated image. The F-score, precision, recall, and accuracy were computed as measures of the agreement between the two. A leave-one-out cross-validation (LOOCV) approach was taken for performance evaluation in order to ensure generalisation of the results to unseen images. The LOOCV results for the various ROCs considered are presented in Figure 2-13.A and summarised in Table 2-7.

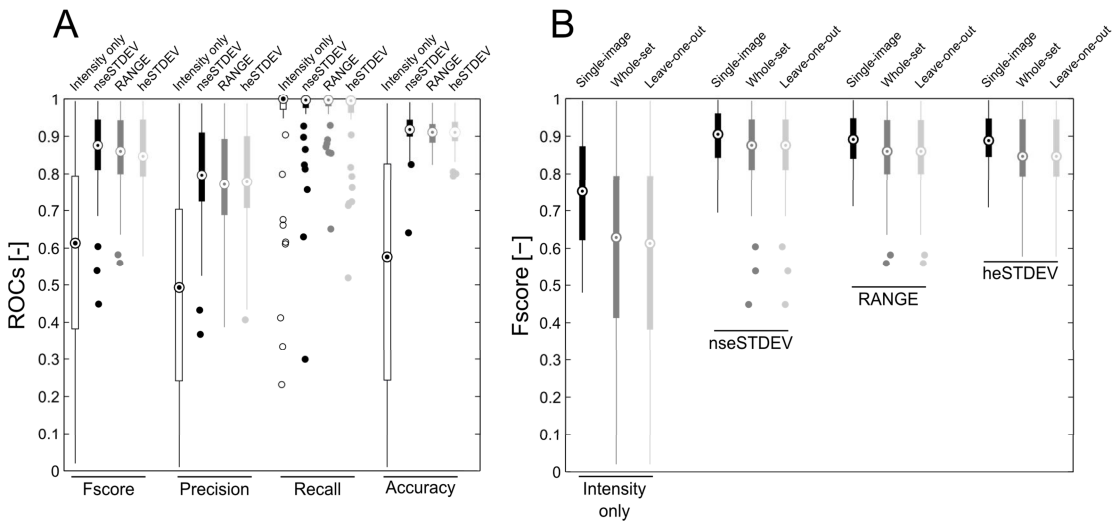


Figure 2-13 Assessment of segmentation performance for three contrast thresholding algorithms and one based on raw intensity thresholding only. (A) Performance metrics for the four algorithms determined using the leave-one-out cross validation method. (B) Impact of various optimisation methods on the mean F-score. In both cases, the circles are the median, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

For all performance metrics considered, the three contrast thresholding algorithms performed significantly better than the one based on raw intensity thresholding. The performance of all three contrast filter-based algorithms was very similar and there was not one approach that stood out significantly from the other two, even though nseSTDEV achieved slightly better median scores in some cases. Intensity thresholding had the broadest distribution of scores, spanning the entire range of possible values for most scoring metrics (i.e. from 0 to 1). This showed its inadequacy for the processing of most of the PCM images considered. Contrast thresholding approaches had much narrower distributions of scores.

Table 2-7 Optimal processing parameters for PCM image segmentation using contrast filters. Leave-one-out cross-validation (LOOCV) was used to asses segmentation performance based on 50 images with F-score as the objective variable (shown as mean ± standard deviation)

	Parameter				Metric			
Algorithms	A	B	C	D	F-score	Precision	Recall	Accuracy
Intensity-only	0.36	1000	25	-	0.590 ± 0.265	0.504 ± 0.286	0.921 ± 0.183	0.543 ± 0.304
RANGE	3	0.05	320	320	0.855 ± 0.103	0.776 ± 0.145	0.972 ± 0.061	0.913 ± 0.043
nseSTDEV	0.6	0.03	320	100	0.856 ± 0.117	0.796 ± 0.141	0.954 ± 0.119	0.915 ± 0.056
heSTDEV	3	0.02	1000	100	0.852 ± 0.105	0.786 ± 0.138	0.951 ± 0.102	0.911 ± 0.048

The ability of the various algorithms to generalise to unseen images was evaluated by comparing segmentation performance determined by single image optimisation (i.e. all parameters are optimised for each image individually), whole-set optimisation (i.e. parameters are optimised for all images in the set at once) and leave-one-out cross-validation (Figure 2-13.B). It was expected that performance would be the highest for the single image method, as

the two other methods involved compromising on the performance for individual images in order to reach an overall optimum. While the median F-score of contrast filters algorithms only marginally dropped between single image optimisation and leave-one-out cross-validation (2.5%, 1.7% and 1.2% decrease for nseSTDEV, RANGE and heSTDEV respectively), the difference for the intensity-only algorithm was 10.4%. Based on these results, contrast thresholding algorithms clearly outperformed raw intensity thresholding.

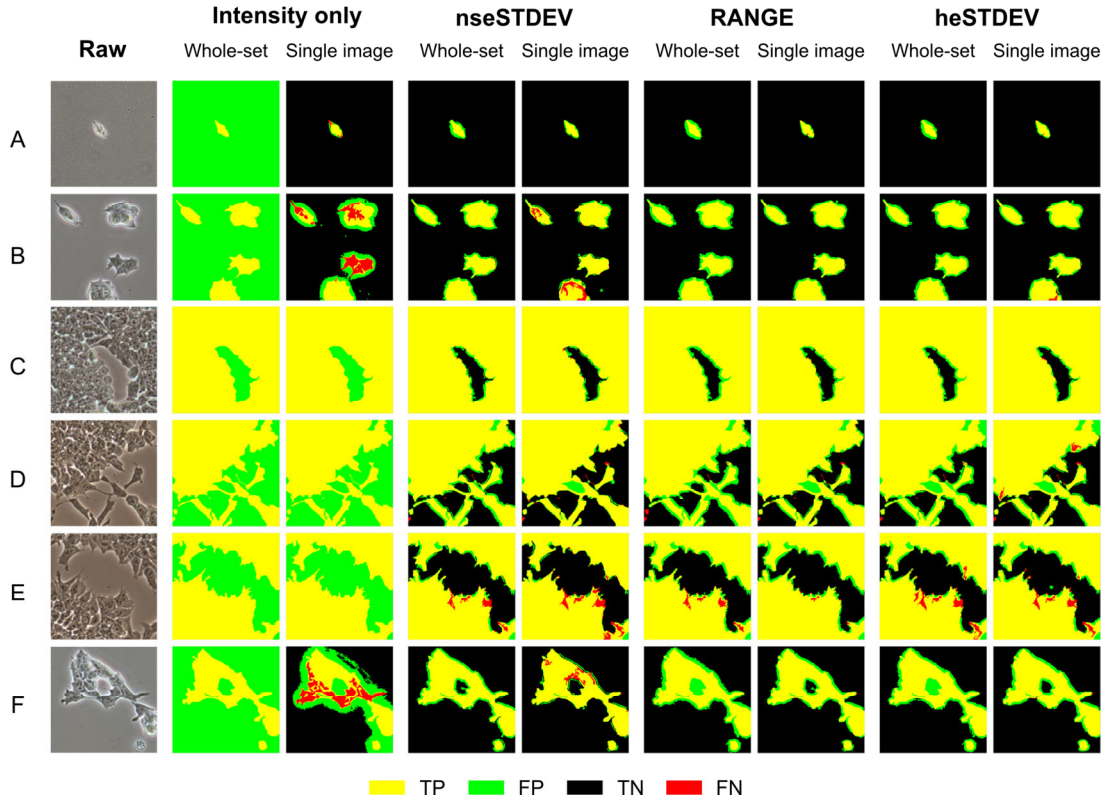


Figure 2-14 Examples of segmentation outputs for the various algorithms tested. Each row is a different image (A-F), the raw PCM image is shown in the first column. For each algorithm, two segmentations outputs are shown: segmentation using parameters optimised for the whole-set (left) or per image (right). These comparison images were obtained by comparing the segmentation output to manually annotated ground truth images. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN).

Closer inspection of the segmentation outputs can be revelatory and clarify what the ROCs obtained translate to in terms of segmentation quality (Figure 2-14). When using parameters optimised for the whole set of images, the algorithm based on raw intensity thresholding only produced images whose pixels were all classified as cells, regardless of the actual content. In contrast, the outputs corresponding to parameters optimised for each image individually showed slight improvements in the segmentation quality. The three algorithms with a contrast-filtering step fared much better and produced segmentation outputs that matched the manually annotated images much more closely. In many cases, objects were surrounded by bands of false positive pixels that corresponded to halo artefacts. When optimising the parameters for each image, the impact of halo artefacts was slightly alleviated as the number of false positive pixels decreased, at the cost of an increase in false negatives. The net effect was still beneficial for some images, indicating that there was room for improvement. These results motivated the development of a

method that could alleviate the impact of the halo artefacts without being detrimental to other aspects of the segmentation.

2.3.2 Improvement of segmentation performance using post-hoc halo correction

Algorithms based on contrast filters were shown to produce good segmentations for the set of PCM images considered. However, all three variants tested were found to be sensitive to halo artefacts. While tweaking of segmentation parameters could help alleviate this effect, it would also result in worsened segmentation performance for other regions of the image. This trade-off was not beneficial for the whole set of images as shown by the optimisation process favouring parameter values that led to misclassification of halo pixels. It was therefore necessary to devise a segmentation approach that would be insensitive to halo artefacts, without compromising the overall segmentation performance, not only in terms of segmentation quality but also in processing speeds.

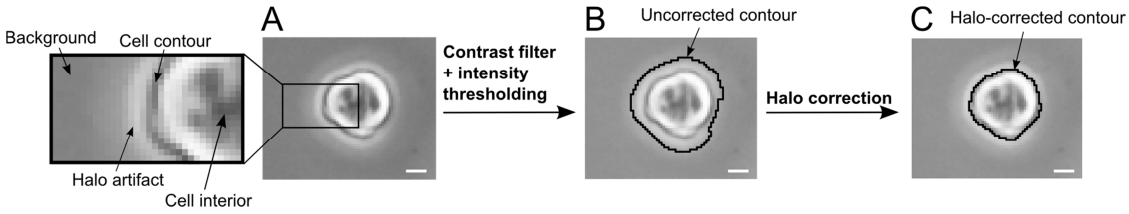


Figure 2-15 Illustration of the segmentation process using contrast filters and *post-hoc* halo correction. (A) Raw PCM image of a single mESC shortly after seeding. (B) Outline showing the segmentation outcome after contrast filtering and intensity thresholding. The pixels on that outline are used to initialise the *post-hoc* halo correction algorithm. (C) Final segmentation result after *post-hoc* halo correction.

It has been previously suggested in the literature that the intensity gradient profiles in the vicinity of cells could be used to detect halo regions (Bradhurst et al., 2008; Seroussi et al., 2012). We thus devised a robust algorithm to harness this property of halo artefact regions so that they could be reliably and rapidly removed from the segmentation output (see method 2.1.4 for details). In short, the gradient direction in the immediate surroundings of cellular objects tended to point inwards (e.g. towards the interior of the cells) as the brightest pixels of halo artefacts were usually found at the interface with the cells. The gradient suddenly changed direction when reaching the cell membrane, whose pixels were much darker than the nearby halo artefacts pixels (Figure 2-7.A and Figure 2-15.A). The proposed *ad-hoc* correction algorithm was initialised at the locations corresponding to the border of objects in the binary output of the segmentation algorithms (Figure 2-15.B). The initial contour was then iteratively shrunk following the direction of the image gradient until it reached the border of the cellular objects where it stopped as the gradient then pointed outwards and the contour was only allowed to shrink, not expand (Figure 2-15.C).

The impact of the *post-hoc* halo correction step on segmentation performance was assessed using a methodology similar to that used for contrast thresholding evaluation. The operations

applied to the input images were kept the same (contrast filtering, followed by thresholding, hole filling, and small object removal). The halo correction was then performed on the binary output produced by these operations. As the optimal parameters for segmentation with halo correction were likely to be different from those obtained previously without correction, the parameter space was screened again. The only additional parameter was one that dictated the maximum fraction of an object that could be shrunk during the halo correction process (Table 2-8). As explained in the method section 2.1.4, this was used to avoid detrimental effects on segmentation quality for rarely occurring PCM images whose features deviated significantly from what was normally expected.

Table 2-8 Explored parameter space for segmentation algorithm (with halo correction).*

		Parameters		
Contrast filter	Operation	Symbol	Name	Screened values
None (Intensity only)	TR	A) ε	Binarisation threshold	0.01 to 1 (step = 0.01)
	PP	B) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		C) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
	HR	D) R_{MAX}	Fraction of object allowed to shrink	{0 ; 0.3 ; 0.7 ; 1}
RANGE	CF	A) W	Window size	11 to 35 (step = 2)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
	HR	E) R_{MAX}	Fraction of object allowed to shrink	{0 ; 0.3 ; 0.7 ; 1}
nseSTDEV	CF	A) σ	Window size	0.1 to 2.5 (step = 0.1)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
	HR	E) R_{MAX}	Fraction of object allowed to shrink	{0 ; 0.3 ; 0.7 ; 1}
heSTDEV	CF	A) W	Window size	1 to 31 (step = 2)
	TR	B) ε	Binarisation threshold	0.01 to 0.2 (step = 0.01)
	PP	C) A_{FOH}	Hole filling area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
		D) A_{RSO}	Small object removal area threshold	{0 ; 25 ; 100 ; 320 ; 1000}
	HR	E) R_{MAX}	Fraction of object allowed to shrink	{0 ; 0.3 ; 0.7 ; 1}

*LC, Local Contrast; STDV, local standard deviation; CF, Contrast filter; TR, thresholding; PP, Post Processing; HR, Halo Removal

The segmentation performance obtained using the additional halo correction step is shown in Figure 2-16.A and summarised in Table 2-9. Algorithms based on contrast filters performed

very well for all ROCs considered, and had narrow scores distributions, which highlighted small variability in segmentation performance for the different training images. The intensity thresholding only algorithm, however, performed poorly and still had distributions of scores spanning nearly the entire range of possible values (i.e. from 0 to 1). These observations were further confirmed by comparison of the segmentation performance as obtained for the various optimisation methods (single image, whole-set, and LOOCV). The F-score for the algorithm based solely on intensity thresholding fell 9.3% between the whole-set optimisation and LOOCV, a slight improvement over the decrease observed without halo correction. In contrast, the decrease for the other algorithms was much smaller at 1%, 0.8% and 0.9% for nseSTDEV, RANGE, and heSTDEV respectively, hinting at a better generalisation to unseen images when the halo correction step was included.

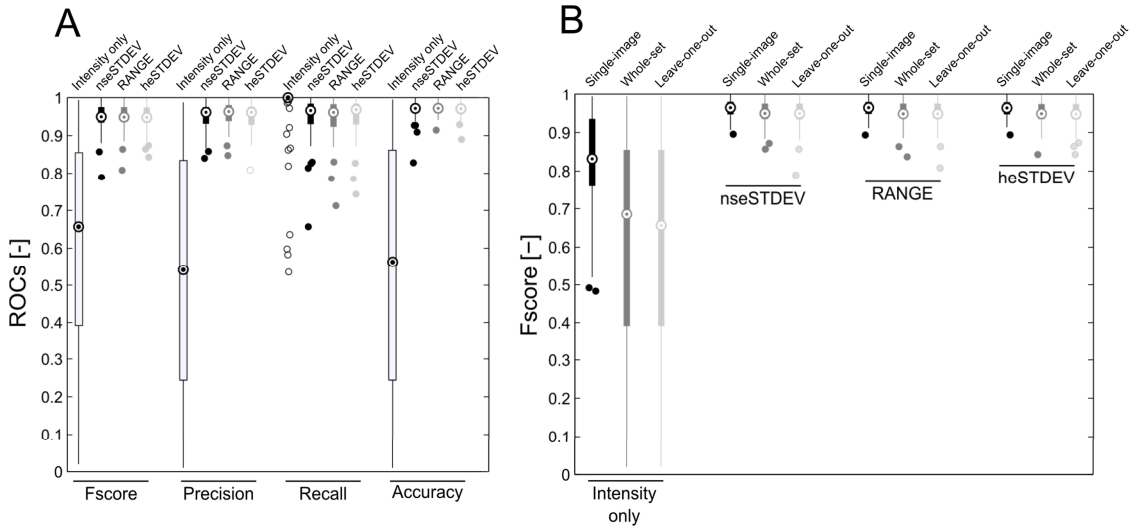


Figure 2-16 Assessment of segmentation performance with halo removal for three contrast thresholding algorithms and one based on raw intensity thresholding only. (A) Performance metrics for the four algorithms determined using the leave-one-out cross validation method. (B) Impact of various optimisation methods on the mean F-score score. In both cases, the circles are the median, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

Similarly to the results obtained without halo correction, the intensity thresholding-only algorithm produced segmentation outputs where all pixels were classified as cells when using parameters as determined based on the whole training set (Figure 2-17). The outputs for parameters optimised on a per-image basis showed a slight improvement but the segmentation was still of poor quality. In contrast, the segmentation outputs for contrast thresholding algorithms were very close to the manual annotations produced by human experts. As such, the outputs of these algorithms were considered high quality. Indeed, key image structures were properly detected, including medium-sized holes within colonies and elongated protuberances. When using parameter values that were determined for each individual image (thus representing ideal performance), small holes in colonies were correctly detected as background. Parameter values obtained by whole-set optimisation, however, led to outputs with some of these holes filled, thus generating false positives. The hole filling post-processing step was based on an area

threshold parameter. The right value for that parameter had to be determined so that holes arising from imperfect segmentation would be filled while legitimate holes (i.e. background space within cellular objects) would be left as is. This task was made challenging by the fact that legitimate holes tended to drastically vary in size depending on the image considered. It is unlikely that this problem can be solved without replacing the current approach with a more advanced technique that would significantly increase the computational complexity of the algorithms. It was therefore accepted that small holes misclassification was a limitation of the algorithms described herein.

Table 2-9 Optimal processing parameters for PCM image segmentation using contrast filters and *post-hoc* halo correction. Segmentation performance was assessed using leave-one-out cross-validation (LOOCV) based on 50 images with F-score as the objective variable (shown as mean ± standard deviation)

	Parameter					Metric			
Algorithms	A	B	C	D	E	F-score	Precision	Recall	Accuracy
Intensity only	0.32	1000	100	0.3	-	0.622 ± 0.268	0.540 ± 0.309	0.953 ± 0.116	0.555 ± 0.316
RANGE	3	0.03	100	320	0.3	0.950 ± 0.035	0.956 ± 0.032	0.946 ± 0.055	0.971 ± 0.017
nseSTDEV	1.1	0.03	320	320	0.3	0.947 ± 0.038	0.951 ± 0.037	0.946 ± 0.062	0.967 ± 0.028
heSTDEV	3	0.01	320	320	0.3	0.947 ± 0.035	0.950 ± 0.038	0.947 ± 0.054	0.968 ± 0.020

For contrast thresholding algorithms, the introduction of the halo correction step led to a significant segmentation score increase for all metrics considered (paired t-tests, p-values < 0.0001) at the exception of recall, which was already very high without correction. The bands of false positive pixels that were previously found around objects due to misclassification of halo pixels were effectively absent after correction. The highly significant increase in performance across all metrics indicated that the ability to correct for halo artefacts was not at the expense of the ability to correctly segment other structures of PCM images.

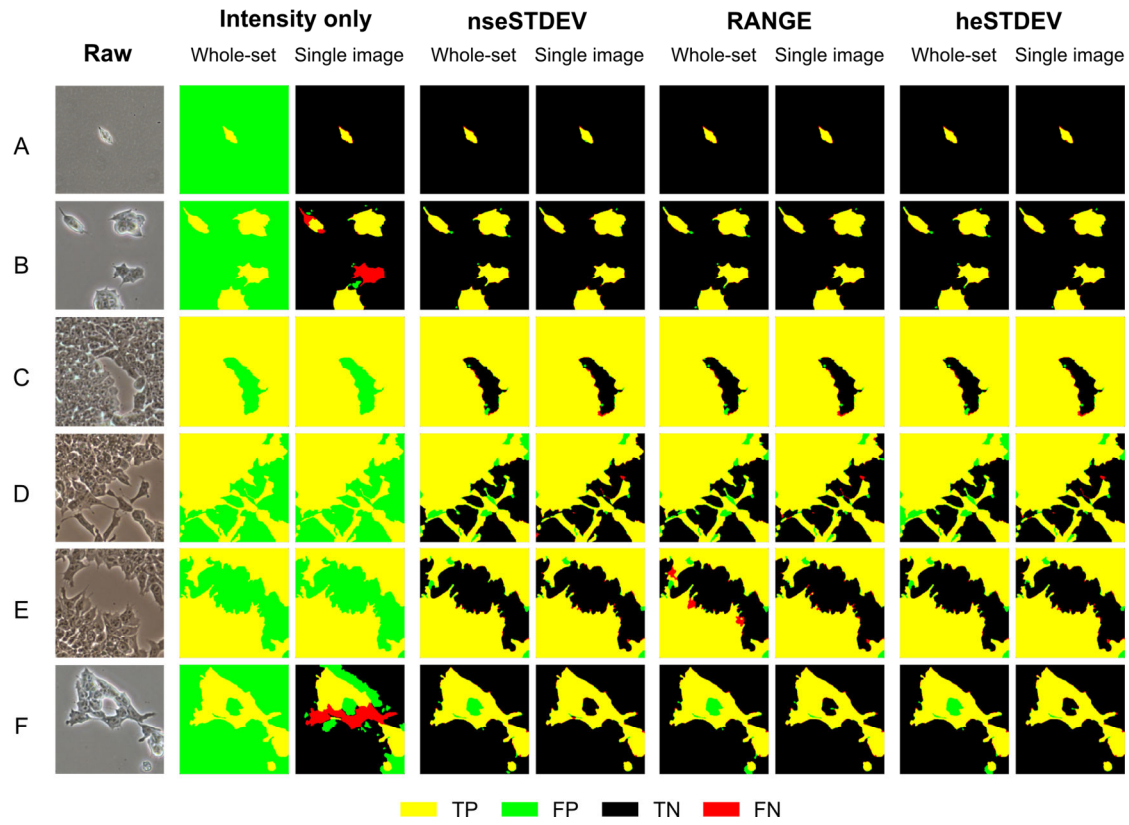


Figure 2-17 Examples of segmentation outputs for the various algorithms tested with halo removal. Each row is a different image (A-F), with the raw PCM image shown in the first column. For each algorithm, two segmentations outputs are shown: segmentation using parameters optimised for the whole-set (left) or per image (right). These images were obtained by comparing the segmentation output to images manually annotated by human experts. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN).

Another interesting aspect of the addition of the halo removal step was that it provided better tolerance to deviations from optimal parameter values. This was illustrated through a simple one-at-a-time "sensitivity analysis"[3] whereby 20 PCM images were segmented using the normalised soft-edged local standard deviation (nseSTDEV) filter for various binarisation threshold values (ε) while all other parameters were kept at their optimal baseline values (as shown in Table 2-7 and Table 2-9). The segmentation performance was evaluated using the F-score normalised to the maximum values obtained for the range of intensity threshold parameter values tested. There was a marked difference between the results obtained for the case without (Figure 2-18.A) and with halo correction (Figure 2-18.B). The decline in performance when deviating from the optimal threshold value was more pronounced and happened for smaller deviations without the halo-correction step. When using halo correction, the performance was found to be very stable for increasing threshold parameter values, while maintaining low variability across the 20 images tested. This was of particular importance as it is often not practical to determine optimal parameters for each set of new images. Having higher tolerance

[3] Sensitivity analysis here is used rather loosely, and simply refers to the characterisation of changes in algorithm outputs as parameter values deviate from their optimal value. It does not imply sensitivity analysis carried out using methods relying on analytical formulations of a problem.

to deviations from optimal values should minimise the loss of performance when a full grid optimisation (e.g. such as the one carried out for this project) is not possible.

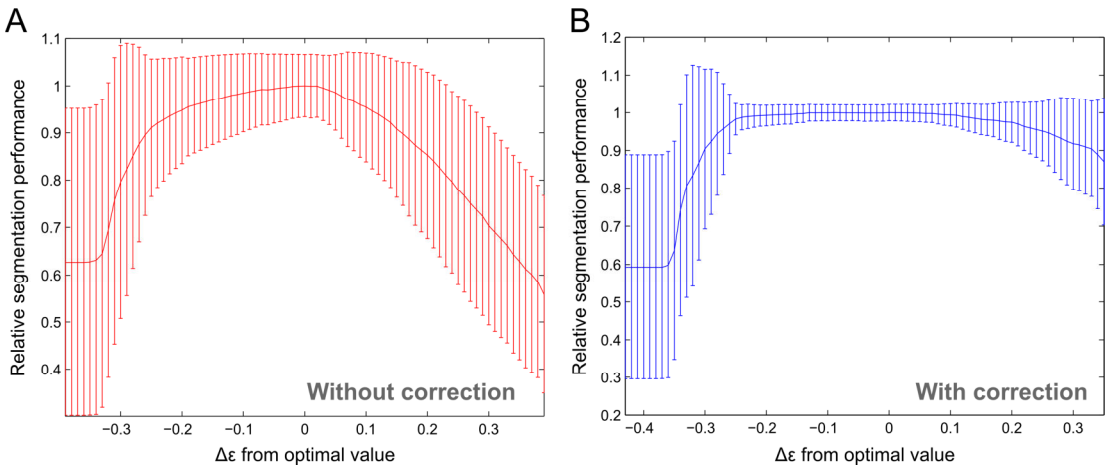


Figure 2-18 Stability of segmentation performance for non-optimal intensity thresholding parameters. F-score for various threshold values (A) without halo correction and (B) with halo correction. The data is shown as mean with error bars representing the standard deviation (n=20 images). The segmentation was based on the normalised soft-edged local standard deviation (nseSTDEV) filter.

Unlike contrast filtering that is mostly based on linear operations, the iterative halo correction algorithm required the use of dynamic programming. Custom code had to be written from the ground up for this particular algorithm. The initial MATLAB implementation of the halo correction algorithm incurred a significant increase in processing time, with the median processing time of a 1280×960 pixels image being about 2.3s, compared to less than 100ms for the contrast filter (nseSTDEV) and intensity thresholding without correction. Re-implementing parts of the halo correction algorithm in C++ led to a significant improvement with a median processing time of less than a second (Figure 2-19.A). Performance could be further improved by taking advantage of parallel computing techniques (trivially implemented in MATLAB using the *parfor* statement).

Given the use of dynamic programming, processing time was partially dependent on the content of an image. This was investigated by measuring the time necessary to process images with varying degrees of confluency (ratio of cell pixels to total number of pixels). Without correction, the time required to segment an image was found to be mostly independent of the image confluency (Figure 2-19.B). In contrast, when using the halo correction algorithm, processing time was clearly impacted by changes in image confluency (Figure 2-19.C). The processing time initially increased together with image confluency, reaching a maximum at around 40% confluency, before linearly decreasing. This was explained by the fact that the number of pixels to visit during the correction process will depend on the number of border pixels initially found after segmentation of the image. This number increased with confluency until large colonies started appearing and filled the field of view, resulting in the linear decrease observed. These results hinted at potential strategies for future optimisation of the algorithm (for example by better selecting which pixels are worth visiting at the next iteration).

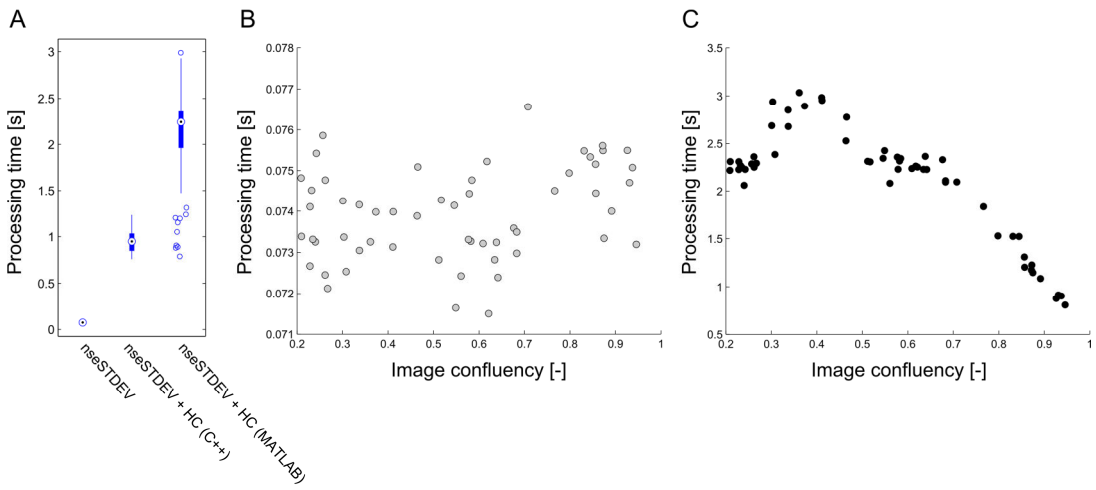


Figure 2-19 Processing time of the iterative halo correction algorithm. (A) Comparison of the segmentation time for PCM images of 1280×960 pixels images. The boxplot shows the results for 56 images. The circles are the median, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)]. Processing time in function of the image confluency (i.e. ratio of cell pixels to the total number of pixels) for the case without (B) and with (C) halo correction.

The combination of contrast filtering and halo-correction produced very high quality segmentation results while maintaining a suitable processing time. The different contrast filters tested (RANGE, hard-edged local standard deviation, and normalised soft-edged local standard deviation) produced very similar results and were mostly inter-changeable. Due to the scope of the project, it was necessary to focus on a single algorithm for further characterisation and application to real-world scenarios. Unless otherwise specified, all the following investigations will be carried out using the nseSTDEV contrast filter with halo removal.

2.3.3 Comparison with other methods

Comparing performance of PCM image segmentation algorithms can be challenging for various reasons. The main issue often relates to the availability of the source-code for said algorithm. Many publications are either based on commercial and closed software or do not provide enough information to ensure a faithful re-implementation of the algorithms for fair comparison. Two methods were found to be suitable for such comparison. The first one was a study by Topman and colleagues that described a very simple yet efficient PCM image segmentation algorithm written in MATLAB (Topman et al., 2011). Their approach was based on the use of multi-scale local standard deviation filters. The other method considered was described by Juneau *et al* in a study that compared a large number of approaches for PCM images processing, with the best performing one being based on a range filter similar to the one evaluated in this work (Juneau et al., 2013). For Topman's approach, the source-code of the algorithm was published alongside the original study and was implemented as is. The algorithm for the method proposed by Juneau *et al*. was implemented using the details provided in the material and method section of the manuscript.

Optimal segmentation parameters and the corresponding segmentation scores were determined for both approaches using the same methodology than the one employed to assess algorithms earlier in this chapter (Figure 2-20). Both Juneau and Topman performed significantly worse than the algorithm based on the nseSTDEV filter with halo correction developed for this work (p-values of 1.7×10^{-12} and 2.8×10^{-11} for Juneau and Topman, respectively, paired t-tests). Juneau used a flood fill algorithm for hole-filling. While efficient, this approach resulted in all holes in the image being filled, regardless of how big they were. For colony-forming cell lines, such as the mESC line used for performance evaluation here, there can be sizeable holes that should be detected as background. Moreover, both methods were sensitive to halo artefacts, further worsening their performance for the set of images considered. Just by adding the halo removal step as described in 2.3.2, the segmentation performance of both methods significantly improved (p-values of 6.9×10^{-10} and 2.1×10^{-12} for Juneau and Topman respectively, paired t-test) and was very similar to that of the approach described in this chapter. These results further confirmed that the halo removal step was critical in achieving the high performance reported earlier in this chapter and that it can be applied regardless of the method used for initial segmentation of the image. Indeed, a one-way ANOVA suggested that the segmentation scores for Juneau, Topman, and the proposed approach (nseSTDEV) were not significantly different when using halo correction (F = 2.52, p-value = 0.084).

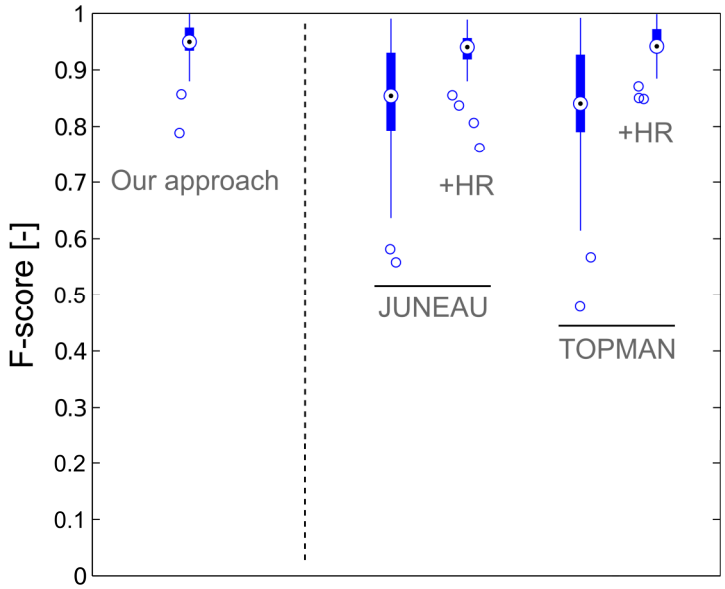


Figure 2-20 Comparison of the proposed segmentation approaches to methods previously described in the literature. Segmentation performance was assessed using leave-one-out cross-validation (LOOCV). +HR denotes that the algorithm had an additional halo removal step applied during processing. The circles are the median F-score for 50 mESC PCM images, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

2.3.4 Generalisation to other cell lines and imaging conditions

The aim of this project was to develop approaches for PCM image processing that could be usable in real-world situations. As such, it would not be satisfactory to limit their validation to a

single cell line, cultured in specific conditions and imaged using a single microscope. In this section, the algorithm is evaluated for PCM images of different cell lines (including non-mammalian cells), acquired using different imaging setups (e.g. microscope models and camera types), and with different imaging conditions (e.g. illumination, focus).

2.3.4.1 Generalisation to other cell lines

Generalisation to other cell lines was assessed in two ways: quantitatively by comparison with ground truth images and qualitatively by visual inspection of segmentation outcomes. First, PCM images of Chinese hamster ovary (CHO) cells and Human neuroblastoma (NB) cells were acquired and processed. CHO cells have traditionally been the organism of choice in mammalian cell-based biotechnology, in particular for the production of therapeutic recombinant proteins (Chu and Robinson, 2001). NB cells have been extensively used as an *in-vitro* model for the study of neurodegenerative disease and neurotoxicity (Cheung et al., 2009). For each cell line, three representative images were chosen and manually processed following the same procedure than that used for the evaluation of performance using mESC PCM images. The segmentation parameters were those determined after optimisation using the set of mESC images (see section 2.3.2). The F-scores for three images was 0.95 ± 0.03 and 0.89 ± 0.08 for CHO and NB PCM images, respectively. The lower score for the neuroblastoma cell images was mostly explained by the presence of very thin, flat and elongated protrusions that are challenging to detect (Figure 2-21). These results indicated that the set of parameters determined using a particular cell line (mESC in this case) was a good starting point for images of other cell lines. Also, a cell-line specific optimisation could likely improve these scores. However, it would require manual annotation of a large number of images, which is tedious and time consuming. It is therefore unlikely that users would wish to go through such a process. This issue will be revisited in section 2.4.2 where rapid determination of optimal segmentation parameters is discussed.

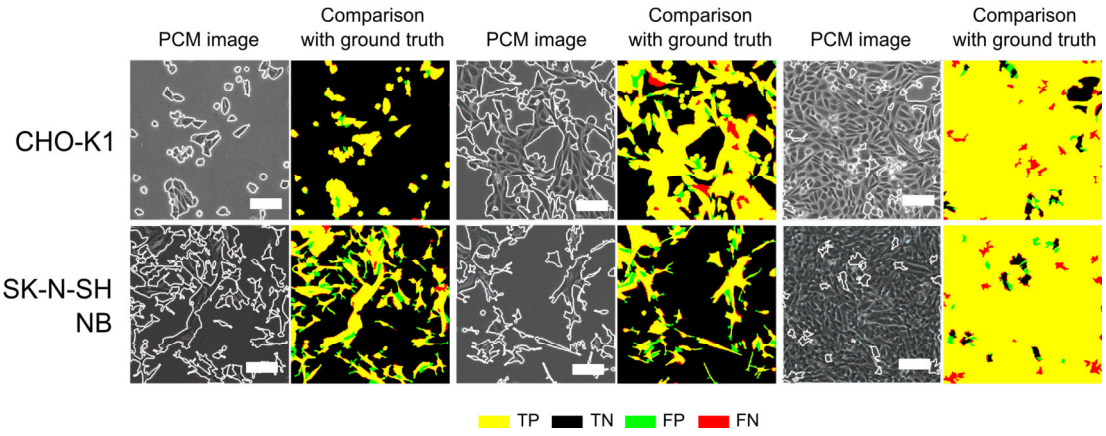


Figure 2-21 Segmentation of Chinese hamster ovary (CHO) cells and Human neuroblastoma (NB) PCM images. The nseSTDEV contrast filter was used, followed by post-hoc halo correction halo correction. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN).

The segmentation performance for a variety of cell types was also assessed qualitatively by visual inspection. The cell types represented in the PCM images considered included 3T3 cells, yeast, Human embryonic stem cells, amoebae and Human sperm cells (Figure 2-22). In most cases, acceptable segmentation could be achieved using the parameters optimised using the mESC images. However, the results shown here were generated using parameters tweaked on a per-image basis to ensure optimal segmentation due to significant differences in image properties (e.g. type of illumination, size of cellular objects) as those were acquired by third parties. These results highlighted the versatility of the algorithm and, by extension, the wide range of potential applications in various areas of biology. It was also clear that it was necessary to provide suitable tools for rapid parameter tweaking when dealing with images whose features differ significantly from those of the training set. A graphical user interface, designed for this purpose, is described in section 2.4.2.

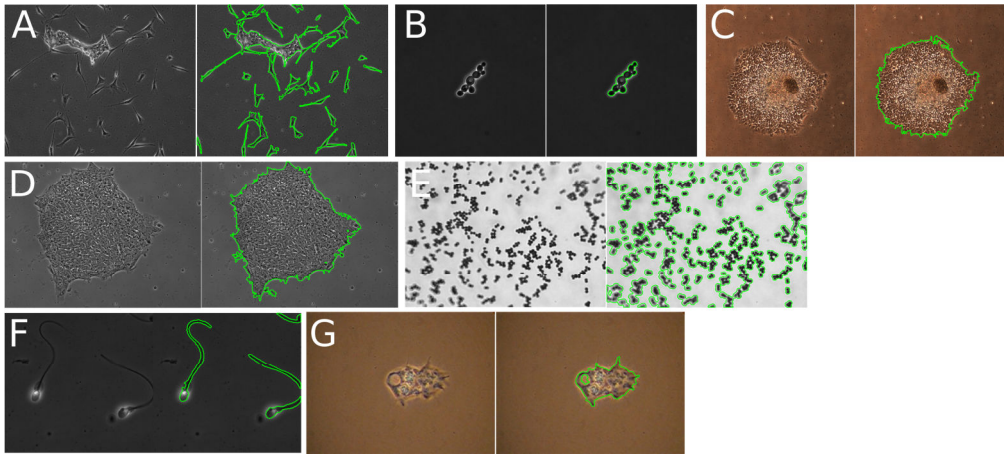


Figure 2-22 Segmentation results for PCM images of a wide range of cell types. For each set, the first image is the raw PCM image and the second is the same raw image overlaid with the segmentation result in green. (A) NIH/3T3, adapted from Cell: An Image Library (John Elliot, CIL 7883). (B) Saccharomyces cerevisiae, adapted from Cell: An Image Library (David Ball, Jean Peccoud, CIL 39626). (C) Human induced pluripotent stem cells cultured on Matrigel, courtesy of Zuming Tang (University College London, Biochemical Engineering). (D) Human embryonic stem cell colony on Matrigel, adapted from Cell: An Image Library (Sabrina Lin, Prue Talbot, CIL 12626). (E) Pichia pastoris, courtesy of Dr Rochelle Aw (Imperial College London, Division of Molecular Biosciences). (F) Human sperm, adapted from Cell: An Image Library (Michael Crammer, CIL 34524). (G) Acanthamoeba castellani, adapted from Cell: An Image Library (Thelma Dunnebacke, CIL 19326).

2.3.4.2 Effect of the imaging setup (microscope model and type of camera)

Imaging setups vary from laboratory to laboratory. Unlike imaging conditions (e.g. illumination, focus, and objective types), the microscope model and camera used for image acquisition are usually fixed and cannot be adjusted on a per-experiment basis in order to optimise image quality to facilitate image processing. The impact of the imaging setup used for acquisition on segmentation performance was assessed by imaging the same cell culture (mESC Oct4-GiP cells grown in 6-well plates) using three different combinations of PCM microscopes and cameras. The combinations were a Nikon Ti-E with a colour camera, an Olympus IX71 with a monochrome camera, and a Zeiss Axiovert 135 with a monochrome camera (more details can be found in section 2.2.2). The distinction between monochrome and colour camera is of

particular importance as colour images were handled slightly differently due to the need of a pre-processing step to merge the three RGB channels into a single grayscale representation. The images were processed using the nseSTDEV algorithm followed by halo removal with optimal parameter values determined earlier in this chapter. The same parameter values were used for all three sets of images. There were not any obvious differences between the segmentation outputs obtained for the images acquired using the three imaging setups (Figure 2-23). This observation was further confirmed by a one-way ANOVA analysis that showed that the imaging setup did not have a significant effect on segmentation performance as assessed using the F-score (df=9, F = 3.75, p-value = 0.09).

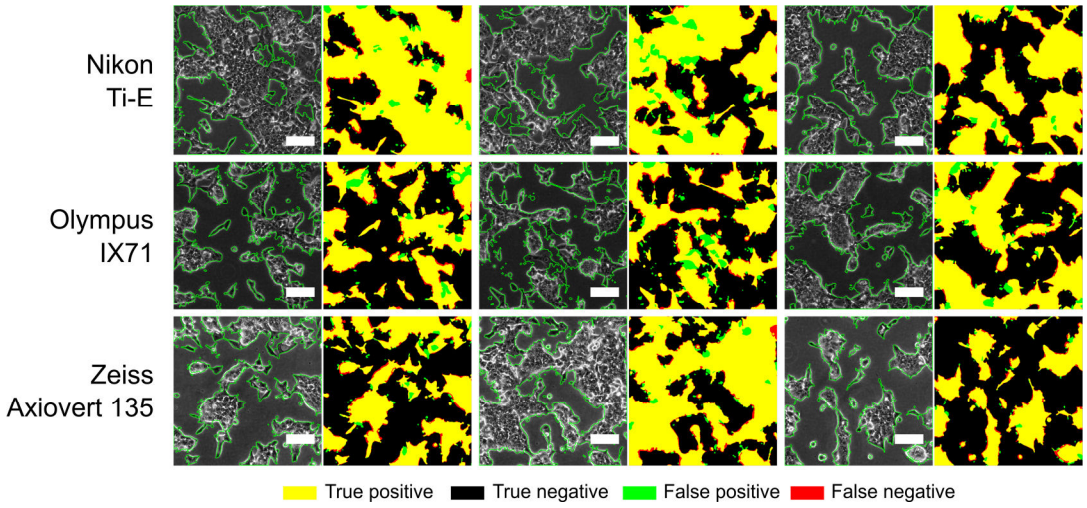


Figure 2-23 Effect of the imaging setup on segmentation outcome. Comparison of segmentation performance for images of a single Oct4-GiP mESC culture acquired using different phase contrast microscopes, cameras, and imaging protocols. The microscopes used were a Nikon Ti-E microscope (Fi-1 color camera), a Olympus IX71 (Hamamatsu ORCA-ER C4742-80-12AG monochrome camera) and a Zeiss Axiovert 135 (Hamamatsu ORCA-R2 C10600-10B-H monochrome camera). Each row is a different microscope. Three fields of view per microscope were considered. For each image, the raw PCM image is shown with the detection overlaid in green on the left, and the comparison with the manually annotated ground truth image on the right. All processing parameters were kept constant. Scale bars are 100 μm.

Due to practical considerations, it was not feasible to image the same region in all three cases. Instead, three random fields of view were imaged for each setup. Similarly, other variables such as illumination intensity, focus level or optical properties (e.g. degree of alignment of the phase plate and phase ring) could not be kept constant across all three experiments. Despite those limitations, the results suggested that a given set of parameter should perform similarly regardless of the imaging setup used if they are at least reasonably comparable.

2.3.4.3 Tolerance to suboptimal imaging conditions

Cell culture laboratories rarely offer optimal imaging conditions. Illumination can be challenging to control due to the lack of shielding from ambient light. Moreover, phase contrast microscopes require re-alignment of the phase plate when imaging new samples, which is rarely done in practice due to time constraints and lack of proper training. Adjusting focus can also be time-consuming, especially when a large number of images are being acquired. Some factors

that are independent of the user expertise can also lead to improper image formation by disrupting the normal function of the phase contrast microscopy method. Such factors include the design of the culture vessel used (e.g. materials that interact with light in detrimental ways) or the formation of a condensation layer on the surface of the vessel due to the difference in temperature between the incubator and the ambient environment. All the aforementioned factors can potentially lead to a deteriorated image quality in real-world acquisition scenarios. It was therefore important to investigate how they impact segmentation performance instead of limiting the characterisation process to ideal training images that are unlikely to be representative of real experimental data.

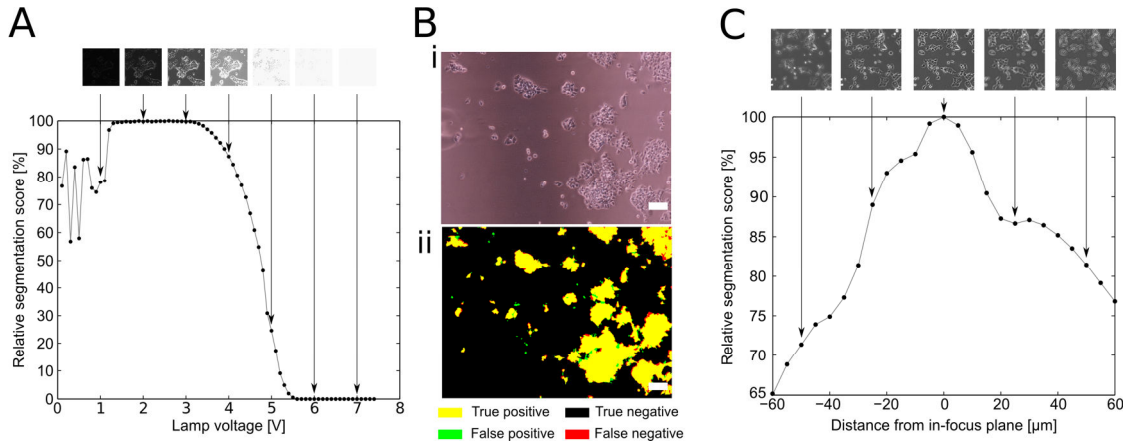


Figure 2-24 Tolerance of the segmentation algorithm to variations in imaging conditions. (A) Effect of the illumination intensity on the segmentation quality (as assessed using F-score). Lamp intensities from 0 to 7 V were tested. (B) Example of segmentation outcome for a PCM image with inhomogeneous illumination patterns as caused by the presence of a liquid meniscus in the light path. Scale bar is 100 µm. (C) Effect of the distance from the in-focus plane on the segmentation quality (as assessed using F-score). In-focus plane was determined visually. The focus was changed in steps of 5 µm using a Piezo Z-stage.

First, the impact of illumination intensity was assessed by acquiring images of the same field of view for lamp voltages ranging from 0V (no direct illumination) to 7V in steps of 100mV (Figure 2-24.A). The lighting of the room housing the microscope was switched off and the microscope area was shielded using a black curtain in order to minimise the effect of ambient light. Despite these precautions, very low amount of ambient light was enough to produce a surprisingly reasonable segmentation, even when no direct illumination was applied. Indeed, the low signal appeared to be amplified by the contrast filter. Nevertheless, the images produced at low voltages (<1 V) were very noisy and variable, as shown by the high score fluctuations recorded. Increasing the illumination intensity beyond 3V became detrimental to segmentation performance as pixels started approaching saturation. Eventually, the whole image was saturated and the structures of cellular objects were lost. The performance was maximal between 1.5V and 3V, which corresponded to voltages that are usually used during normal operation of the microscope. It was therefore not critical to precisely adjust illumination as long as the intensity was reasonable, and small variations that can be expected from an uncontrolled environment (e.g. microscope lamp, ambient illumination) did not significantly impact segmentation performance.

Inhomogeneous illumination patterns are common occurrences when imaging live cells in culture vessels using phase contrast microscopy. Examples of such patterns are those forming due to the presence of a liquid meniscus, a curved air-liquid interface caused by surface tension. The curvature results in light refraction, thus interfering with the phase contrast image formation mechanisms (see section 1.2.2). This is mostly noticeable for small-scale culture vessels such as 96-well plates or at the edges of larger wells. The impact on image quality can be severe and often precludes the detection of the cells using a simple intensity threshold (Figure 2-24.B.i). However, using the nseSTDEV contrast filter algorithm followed by halo correction enabled good segmentation performance, using the same optimal parameters than those determined for mESC images unaffected by such artefacts (Figure 2-24.B.ii).

Focusing is another fundamental aspect of microscopy that could potentially lead to degradation of image quality if done improperly. The seemingly obvious concept of an in-focus image can be difficult to formulate, especially when dealing with complex cellular objects. Experienced users are usually able to rapidly attain a reasonable focus. However, whereas it would only slightly affect perception by humans, small deviations from optimal focus can result in poor segmentation performance due to drastic changes in the image at the pixel level. This was shown by imaging a single field of view for z-positions in increments of 5µm. The results showed that a deviation of only 10µm from the optimal in-focus plane (as determined by experienced microscopy users) could lead to a decrease of 5% in relative segmentation performance (Figure 2-24.C). It is therefore critical to ensure good focusing during image acquisition. This issue can be minimised by using a microscope that is equipped with automated focusing (e.g. Nikon's Perfect Focus System), which enables the acquisition of multiple images at different locations of a culture without having to adjust the focus for individual fields of view.

2.4 Open-source implementation: PHANTAST

One of the aims for this project was the development of a high performance method for the monitoring of adherent cell culture that could be used by researchers of non-technical backgrounds in real-world situations. Users should not be expected to be able to program or use complicated scripting languages. Instead, a simple and clear graphical user interface (GUI) should enable anyone to generate high quality data in a short amount of time. The time to first result (TTFR) is critical: a new user should get the first results in just a few minutes, which requires the software to be designed in a way so that tedious configuration and parameter tweaking can be avoided altogether. These are the reasons why all the algorithms developed for this project were compiled in the form of the phase contrast segmentation toolbox (PHANTAST), which was released under a permissive open-source license. Moreover, graphical tools were built specifically to help non-technical users generate data as quickly and easily as possible.

PHANTAST was released in the following forms:

- MATLAB and C++ source-code (requires a MATLAB license)
- Standalone GUI based on MATLAB (require the installation of free MATLAB runtimes)
- ImageJ/FIJI plugin (entirely free)

A homepage was built for the distribution of the code and that of the various binaries[4].

2.4.1 PHANTAST GUI for adherent culture PCM image processing

The graphical user interface (GUI) of PHANTAST was built to simplify the processing of PCM images as much as possible. It facilitated the determination of optimal segmentation parameters and enabled rapid processing of large number of images for the determination of useful characteristics of adherent cell cultures (as shown in chapter 3).

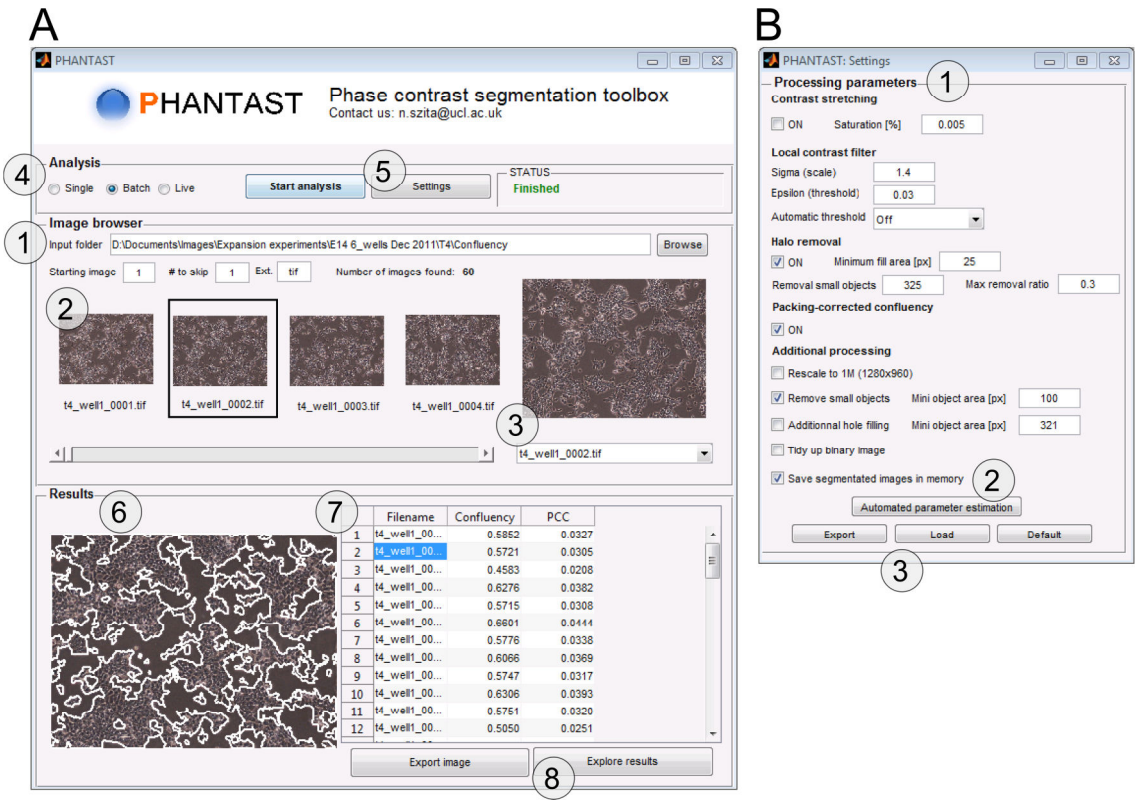


Figure 2-25 Screenshots of the PHANTAST graphical user interface (GUI). (A) Main window of the GUI. (B) Processing settings window.

The PHANTAST GUI is based around the concept of an input folder containing the images to process (Figure 2-25.A.1). This approach was chosen, as it was the most straightforward way of having the user input a large number of images if batch processing was necessary. Images from different experiments could be stored in their respective folders and analysed all at once. The

[4] http://tinyurl.com/phantast or http://code.google.com/p/phantast/

images located in the input folder could be visualised in an image browser (Figure 2-25.A.2). The browser also enabled the selection of images in the case where the processing is done for a single image at a time. In addition, a dropdown list could be used for the selection of the image of interest based on its filename (Figure 2-25.A.3).

There were three image processing methods implemented in the GUI (Figure 2-25.A.4): single image (analysed the selected image only), batch (analysed all images present in the input folder) and live (used to optimise segmentation parameters, as discussed in section 2.4.2). Segmentation parameters could be tweaked using the settings window (Figure 2-25.A.5 and Figure 2-25.B.1). When using the 'live' analysis option, the changes in parameters were immediately applied, allowing empirical determination of optimal parameters. In addition to manual tweaking, an automated parameter estimation tool could also be used to determine optimal segmentation parameters with minimum user input (Figure 2-25.B.2 and section 2.4.2). Parameter values could then be saved and loaded directly from the settings window, enabling users to generate a collection of optimised parameters for different image sets (Figure 2-25.B.3).

During processing, the status box was updated with the percentage of completion and the preview window displayed the segmentation output for the most recently processed image (Figure 2-25.A.6). It was possible to display any image of the user interface in full screen for inspection by simply clicking on the corresponding thumbnail. The actual numerical processing results were presented in a table (Figure 2-25.A.7), which could then be copy-pasted into other software packages for further analysis (e.g. Excel, MATLAB or R). In addition, it was possible to visually inspect the segmentation results for all the processed images in a full-screen window by using the explore results functionality (Figure 2-25.A.8).

In practice, this GUI proved to be very effective. Most of the results presented in chapter 3 were generated using this tool. Other users were able to quickly adopt it and generate quantitative results for their adherent cell culture PCM images.

2.4.2 A tool for automated parameter estimation based on sparse annotations

For the validation of the algorithm, it was necessary to perform an exhaustive grid search in order to determine the optimal segmentation parameters for a given set of images. This required the manual annotation of ground truth images (i.e. labelling every pixel as either cell or background), which is time-consuming and error prone. Depending on the complexity of its content, even a small image of size 250×250 pixels can take up to 20 minutes to properly annotate. Given that a large number of images is preferable for parameter optimisation, the users cannot be expected (and most likely will not) spend such an impractical amount of time for parameter optimisation.

Alternatively, parameter optimisation can be done by 'knob turning', which consists in manually adjusting parameter values and visually assess the quality of the segmentation. This approach can be quick but is qualitative and subjective, as such the resulting parameter values are unlikely to be optimal or generalise to unseen images. The lack of rapid and quantitative approach for segmentation parameter optimisation is often thought to be a barrier to the adoption of automated image processing methods by potential users (Pretorius et al., 2011). It was therefore necessary to devise a method that alleviates this issue.

Trainable and interactive segmentation approaches are getting more and more popular as they allow users to obtain high segmentation performance based on inputs that take the form of paint strokes (Sommer et al., 2011a). These methods are based on machine-learning classifiers, which are well suited for this type of sparse training input whereas traditional approaches tend to rely on fully annotated images. In addition, the training (or off-line) stage for these methods do not usually require computationally expensive grid-search optimisations but instead rely on more elaborate learning techniques that are often much more efficient and less time-consuming.

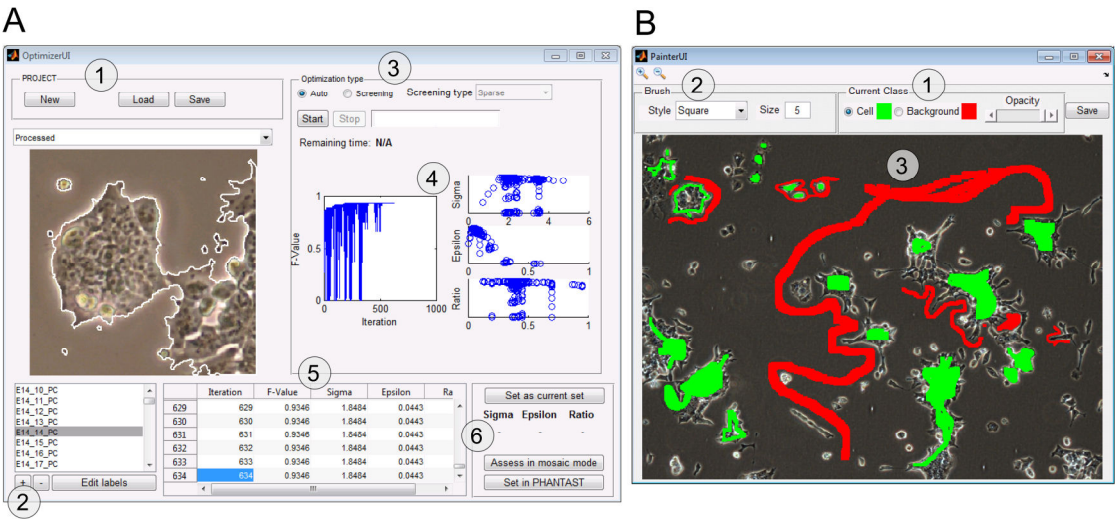


Figure 2-26 Screenshot of the optimisation tool. (A) Main window and (B) labelling window.

We set out to develop a tool that would enable the convenience of sparse annotations (i.e. using a simple painting tool to annotate only a fraction of an image) combined with a pattern search optimisation approach that drastically reduced the time necessary to obtain reasonably optimal values for the segmentation parameters. The operation of the optimisation tool was based on the concept of projects. Each project contained a set of images, their respective annotations, and optimisation results. Projects could be created, saved, and loaded directly from the GUI, enabling the storage of optimisation results for later analysis (Figure 2-26.A.1). Images added to a project were manually annotated using the painter tool (Figure 2-26.A.2 and Figure 2-26.B). This tool showed a full-screen representation of the current image and allowed to label pixels as either cell or background using different brushes (e.g. square, rounded) of various sizes (Figure

2-26.B.1 and Figure 2-26.B.2). This was done by drawing directly on top of the image like in any other image editing software (Figure 2-26.B.3).

After annotation, two optimisation modes could be used (Figure 2-26.A.3): automatic and screening. The former used a pattern search optimisation algorithm to determine the optimal parameter values whereas the latter was a grid search of various levels coarseness (e.g. sparse for small grids to dense for large grids). The automatic method was usually preferred as it enabled to rapidly (in less than 10 minutes) determine optimal parameters without having to explore a pre-defined parameter search grid. During optimisation, various plots enabled the visualisation of the optimisation progress (Figure 2-26.A.4). All results were stored in a table that could be copied and pasted into other software packages for further analysis such as Excel, MATLAB or R (Figure 2-26.A.5). When satisfied with the optimisation outcome, the user could decide to save the parameter values obtained and send them to the main PHANTAST GUI in order to start processing images.

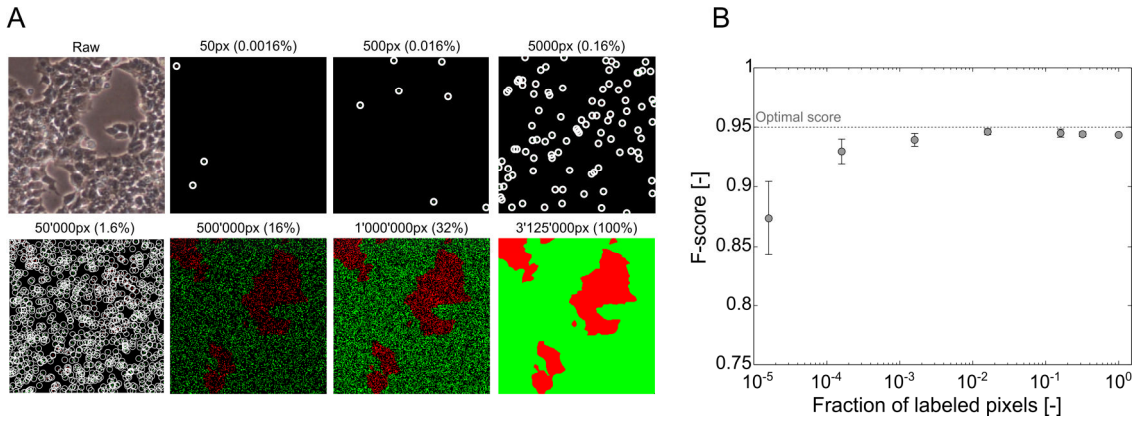


Figure 2-27 Parameter value optimisation based on sparse image annotation.(A) Synthetic sparse annotations generated from a fully annotated image. The number of pixels sampled across the whole input set (50 images) is shown above each image. Pixels in green are cell annotations; pixels in red are background annotation. For clarity, pixels were circled for the first four images. (B) Performance of the optimal segmentation parameter values as the fraction of labelled pixels (logarithmic scale) was varied from $1.6*10^{-3}$% (50 pixels) to 100% (3125000 pixels). 50 fully annotated PCM images were used for this study. The data points represent the mean of the scores obtained for three (n=3) random pixel sampling and the error bars are the standard deviation. The dashed-line represent the optimal score obtained through extensive grid-search using the 50 fully annotated images.

When the optimisation mode was set to automatic, the tool used the Nelder-Mead (or downhill simplex) optimisation method, which enables the minimisation of functions whose derivative is not known (Nelder and Mead, 1965). The choice of this method was arbitrary and was mostly dictated by its availability in MATLAB's optimization toolbox as well as the performance obtained during preliminary experiments comparing it to other methods (e.g. genetic algorithms, particle swarm, simulated annealing). It is likely that Nelder-Mead is not the best choice for this optimization problem but the characterisation of said problem (e.g. smoothness, convexity) and the subsequent selection of an optimization methodology was out of scope of this work as this was only used as a proof of concept to illustrate automated parameter determination.

The ability of this method to determine optimal parameter values based on sparse image annotations was assessed by comparing the segmentation performance using said values to that of the values determined using an extensive and time-consuming grid search as reported in section 2.3.2. The previously mentioned set of 50 fully annotated mESC PCM images (see section 2.2.3) was used as a starting point. Sparse annotations by a user were simulated by randomly sampling varying number of pixels across the 50 images (Figure 2-27.A). The Nelder-Mead method was then used to minimise the segmentation error (e.g. 1-F-score) computed based on the sampled pixels. Due to the random nature of the sampling, this was repeated three times for each fraction of image pixels considered. Segmentation performance of the resulting parameter values was then evaluated using all the pixels of the 50 images of the training set (Figure 2-27.B). As expected, higher fraction of labelled pixels led to parameter values resulting in higher segmentation performance, with smaller variations due to the random sampling. Using as little as 0.01% of the pixels available (representing 31250 pixels in total), it was possible to achieve performance that was very close to that obtained using the values from the extensive grid search.

This approach not only allowed the generation of near-optimal parameter values in a fraction of the time required for an extensive grid-search, it also enabled to do so by annotating only a small fraction of the input images pixels. The time to first result (TTFR) was thus greatly shortened.

Further characterisation of this method is necessary. For this study, the sampling was done irrespectively of the number of cell and background pixel, therefore creating a bias towards one or the other. Ideally, an equal number of pixels would be sampled for each class. In addition, random sampling of the pixels is not necessarily a good representation of user-made annotations. Indeed, these annotations would likely take the form of long strokes instead of sparsely distributed dots. It would be interesting to devise an approach to simulate such annotations and see how they impact the optimisation process. Here, pixels were sampled from 50 different images, which is still a large number of images. It would also be of interest to determine how the number of sparsely annotated images would affect the generalisation of the parameter values obtained after optimisation.

2.4.3 ImageJ/Fiji plugin

The open-source release of the MATLAB/C++ code was a first step to make the algorithms developed during the course of this project available to as many researchers as possible. However, it still required the use of proprietary software (MATLAB) or at least the installation of the corresponding (freely available) runtimes. In order to offer a completely free and open alternative, a plugin was developed for Fiji (Fiji Is Just ImageJ), a distribution of the

ubiquitously used biomedical image processing software package ImageJ (Schneider et al., 2012). This required the port of the MATLAB/C++ code to Java as well as the replacement of MATLAB built-in functions by either the corresponding functions in ImageJ or by ones written from the ground-up for this work. With the exception of a few edge cases where different implementations of built-in functions resulted in slight variations, the Fiji plugin produced results that were identical to those generated using MATLAB (Figure 2-28.A). The plugin was designed to be compatible with ImageJ's macro language, which enabled the emulation of the batch-processing mode that was implemented in the MATLAB version of the GUI (Figure 2-28.B).

Not only this plugin enabled virtually any researcher to use PHANTAST, it also facilitated its integration with the very large library of algorithms and plugins available for ImageJ and Fiji. PHANTAST could also in the future be ported to other image processing packages, such as Cell Profiler (Carpenter et al., 2006).

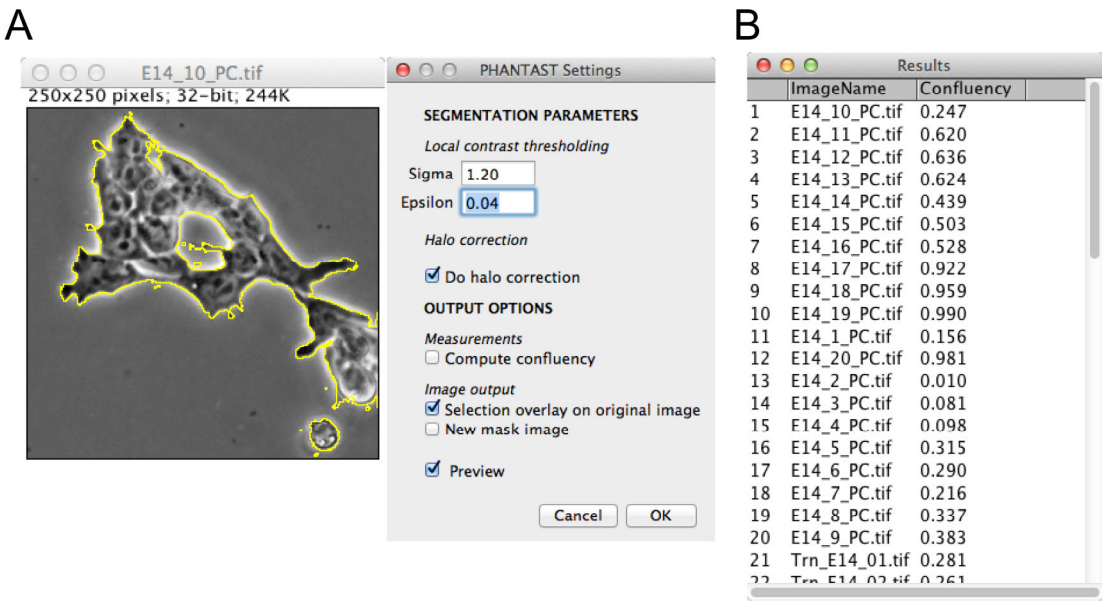


Figure 2-28 ImageJ/Fiji implementation of PHANTAST. (A) Main window showing an example of segmentation using the preview function. (B) Example output of batch determination of confluency using FIJI's macro function

2.5 Conclusion and outlook

A novel algorithm for the segmentation of adherent cell on PCM images was successfully designed, implemented, and validated. The proposed algorithm was based on the concept of contrast filtering, which greatly enhanced the contrast between cell and background, enabling the detection of cellular regions by simple thresholding. Dynamic programming was then used to correct for halo artefacts so that actual contours of the cellular objects could be detected. This approach resulted in very high segmentation performance and compared favourably to previously described methods. The algorithm was designed to have a low computational

complexity, which allowed maintaining processing time under a second for standard resolution PCM images (1280×960 pixels or ~1.2 megapixels).

The algorithm generalised well to other cell lines, demonstrating a high degree of versatility. It also coped well with variations in imaging conditions (e.g. illumination intensity), which made it suitable for use in a cell culture laboratory where the environment is rarely optimal for image acquisition. While it was found to perform well in most situations, edge cases that break some of the assumptions made might arise. The most likely example would be instances where the transition from halo regions to the cell contours does not produce the expected inversion in gradient. This might happen directly after mitosis when the cells are as bright, or brighter, than the surrounding halos or when either the halo or cell boundaries are broken and discontinuous (Ali et al., 2011; Bradhurst et al., 2008). In those cases, the *post-hoc* halo correction would continue until the gradient inversion eventually occurs due to intra-cellular structures. But at this point, a large portion of the cellular objects would likely have been misclassified as background. This is in part alleviated by the R_{max} parameter which controlled the maximum fraction of an object that could be shrunk by the halo correction algorithm. An alternative would be the use of active contour models that are constrained by forces so that they maintain their shape regardless of the presence of irregularities in the halos or cell boundaries (Ambühl et al., 2012; Seroussi et al., 2012). These approaches were evaluated but resulted in a significant impact on processing time, which was not deemed a worthwhile trade-off considering the low frequency of such issues arising in the images considered. It would certainly be of interest to develop a method that retains the simplicity of the gradient-steered shrinking algorithm used in this work while at the same time incorporating some of the fundamental functionalities and properties of active contour models (e.g. shape constraints).

The algorithm presented here, much like a majority of the previously described approaches, relies on *a priori* knowledge of the properties of PCM images, such as the pixel intensity homogeneity differences between cell and background regions, or the presence of halo artefacts. Based on this knowledge, a series of operations were hand-crafted so that the desired output would be produced by the algorithm. This approach to image processing problems in general, and that of segmentation in particular, has recently been challenged by methods that employ machine learning techniques to avoid the painstaking development and optimisation of specialized filters and operations by instead relying on statistical classifiers to learn how to best process images based on their features. In principle, this would allow the development of a very generic approach whereby new inputs and user-defined hard constraints are learned by the machine learning classifiers without requiring any problem-specific knowledge. There is a case to be made that such approaches rely on poorly understood black boxes, yet they appear to be very effective and are found to often outperform specialised algorithms. This point will be further discussed in chapter 4 where the contrast thresholding and halo correction algorithms

will be compared to a more general and trainable segmentation method based on machine learning.

All the various components described in this chapter were embedded in a framework termed the phase contrast microscopy segmentation toolbox, or PHANTAST. Most importantly, PHANTAST included a graphical user interface (GUI), which enabled rapid batch processing of large number of images. The GUI also included a drawing-based tool for the rapid determination of optimal segmentation parameters without any prior knowledge of image processing concepts. Such a tool was designed to drastically reduce the barrier of entry for new users and decrease the time to first result (TTFR). PHANTAST has been made available under a permissive open source, which allows any use (including commercial) as long as the work is properly attributed. This will enable the code to be further improved or integrated in other software pipelines. The algorithm was also made available as a plugin for ImageJ/Fiji, a software package for image processing that is widely used in the fields of life sciences.

Whereas this chapter focused on the theoretical aspects of the algorithm and on its validation, the next chapter shows how PHANTAST can be used to monitor mouse embryonic stem cell cultures in various relevant scenarios.

Chapter 3

Imaging-based monitoring of adherent cell cultures

The cytotoxicity experiments described in this chapter (section 3.2.3) were carried out by Ana Keser, a visiting student from the "University of Applied Sciences, Western Switzerland", as part of her Bachelor thesis diploma work. Ana maintained the cells, performed the experiments, and acquired the microscopy images that were used for the analysis presented here-in.

Chapter 2 described approaches for the segmentation of cellular objects on phase contrast microscopy (PCM) images. The algorithm retained after thorough evaluation combined a normalised soft-edged standard deviation (nseSTDEV) filter with a *post-hoc* halo correction performed using dynamic programming techniques. This algorithm was shown to produce high quality segmentation, i.e. accurate classification of pixels as either cell or background. It was also quick, generalised well to a wide range of cell lines, and was robust to variations in image acquisition conditions.

In this chapter, the segmentation algorithm will be applied to generate useful, quantitative measurements of adherent cell culture characteristics in various relevant experimental scenarios.

3.1 Deriving informative data from PCM image segmentation output

Microscopy methods are routinely used for visual inspection of cell culture devices. In fact, visual inspection often constitutes the only mean to assess the state of a culture non-invasively during its course. The generation of this information relies on human vision and our ability to interpret complex visual stimuli. Indeed, trained operators can rapidly assess the state of a culture or evaluate its confluency. However, these measurements are both qualitative and operator-dependent. As such, they are often unsuitable for experimental outcome assessment and decision-making, potentially resulting in poor reproducibility across trials and laboratories.

This section introduces a palette of methods that can be used to obtain unbiased and quantitative information based on the result of PCM image segmentation: confluency estimation, morphometric analysis, and augmented fluorescence image generation. If not otherwise

specified, the results shown were generated using the normalised soft-edged local standard deviation filter followed by halo correction, using the optimal parameters determined by extensive grid search (Table 2-9).

3.1.1 Confluency estimation

The size of a cell population is a key characteristic of adherent cell cultures. However, enumeration of individual adherent cells using light microscopy methods is often unfeasible due to ill-defined boundaries between cells (e.g. cells growing in colonies, see Figure 2-1 in Chapter 2 for an example) or the low contrast between cells and the background (e.g. flat cells). Even in cases where cells remain as well-delimited individual objects (e.g. immortalized cell lines such as Chinese hamster ovary cells), enumeration is tedious and highly time-consuming. Instead, confluency (or confluence) is often used as an alternative. It is loosely defined as the fraction of the growth area occupied by cells. Confluency has been widely adopted due to the fact it can be determined easily and rapidly by visual inspection using commonly available microscopes. It is especially useful for decision-making when detachment is not possible, for example to determine the optimal time for culture passaging (Kato et al., 2010) or for a perturbation (Dong et al., 1998; Jeong et al., 2005; Stroka et al., 2012). Unlike cell counting, confluency also informs on the spatial organisation of the cells, a property that can affect cell function. Indeed, the degree of confluency of a culture was shown to impact gene expression (Ruutu et al., 2004), the formation and maintenance of intracellular junctions (Lampugnani et al., 1997), and the ability of embryonic stem cells to give rise to viable embryos (Gao et al., 2003).

As such, confluency may become the de-facto standard for non-invasive adherent cell culture monitoring, possibly also due to the lack of a viable alternative. However, it was previously shown to be highly operator-dependent (Topman et al., 2011). It was therefore of interest to determine whether the developed image processing algorithm could improve the robustness and reliability of confluency measurements.

3.1.1.1 Investigating variability in human operator confluency estimation

First, the variability in confluency estimation by human operators was investigated by asking experienced cell culture researchers to estimate culture confluency from sets of mouse embryonic stem cell (mESC) PCM images to mimic routine inspection of a culture. The aim was to determine inter and intra-variability between researchers, or in other words, how estimates from different researchers compared and how consistent a researcher was with him or herself.

At six time point of a culture in 6-well plates (3, 23, 30, 48, 53, and 81 hours after seeding), six images were acquired at random locations of a well (Figure 3-1.A). For each time point,

experienced researchers (> 1 year of cell culture and microscopy experience) were tasked with estimating confluency based on the six images of the corresponding set. The different sets were presented in non-chronological order. Intra-variability was tested for with an additional set of images, which was identical to the 30 hour set but had its images flipped, rotated, and re-ordered. In principle, researchers should produce identical estimates for these two sets, as the images they contained were identical.

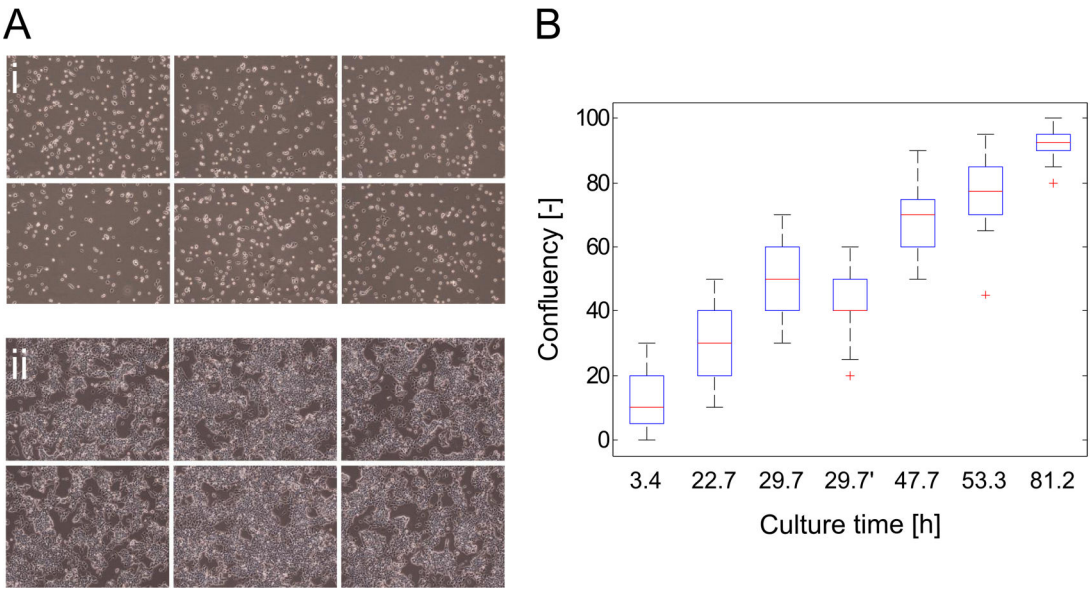


Figure 3-1 Evaluation of confluency determination variability. (A) Examples of sets of mESC PCM images given to experienced researchers. For each set, 6 PCM images were acquired at random location (i) 5.8h and (ii) 71.3h into a culture in 6-well plates. (B) Boxplot summarising the confluency estimations by experienced researchers. For each time point, a set of 6 PCM images was used as a basis for estimation. 6 time points were considered and the sets were presented in non-chronological order. In addition, an extra set of image was included, which contained the same image as the 30h time point set but re-ordered and flipped. The red bars are the median confluency estimation for 14 experienced researchers, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The '+' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

Given m = 7 set of 6 images, n = 14 researchers, $e_{i,j}$ the j-th person's estimate of the i-th set of images, and $\overline{e_i}$ the mean estimate for the i-th set of images, the combined variability of the estimation can be computed as:

$$V = \sqrt{m^{-1}\sum_{i=1}^{m}\frac{\sum_{j=1}^{n}\left(e_{i,j}-\overline{e}_i\right)^2}{(n-1)}} \qquad 3.1$$

The intra-variability can be computed from the repeated estimates $r_{\alpha,1}$ and $r_{\alpha,2}$ by the n researchers based on the two identical image sets:

$$V_{\text{intra}} = \sqrt{\frac{\sum_{\alpha=1}^{n}\frac{1}{2}\left(r_{\alpha,1}-r_{\alpha,2}\right)^2}{n}} \qquad 3.2$$

Finally, the inter-variability can be determined from the combined and intra-variability values:

$$V_{\text{inter}} = \sqrt{V^2 - V_{\text{intra}}^2} \quad 3.3$$

The results showed very high variability in confluency estimation (Figure 3-1.B). The inter- and intra-researcher variability were found to be 9.5% and 6.8% respectively, resulting in a combined variability (i.e. the precision of the researchers' estimation) of 11.7%. Quantifying this variability was essential to determine the extent to which automated culture confluency estimation is beneficial.

3.1.1.2 Automated image confluency estimation

Image confluency can be defined as the ratio of pixels classified as cells to the total number of pixels. This ratio can be computed directly from the binary output of the segmentation step. The quality of the segmentation will thus directly dictate that of the confluency estimation. At this stage, it is important to note that image confluency is not equivalent to culture confluency due to sampling considerations that will be investigated later (section 3.1.1.3).

The quality of the image confluency estimation was assessed using various metrics. First, given $\hat{x}_i$ the estimated image confluency from the output of the segmentation algorithm and x_i the corresponding confluency value determined from the manually annotated ground truth image, the root mean square error for *n* images was computed as follows:

$$RMSE = \sqrt{n^{-1}\sum_{i=1}^{n}\left(\hat{x}_i - x_i\right)^2} \quad 3.4$$

The bias (or mean signed difference), which informed on the systematic difference between the mean of repeated measurements and the true value, was given by:

$$bias = n^{-1}\sum_{i=1}^{n}\left(\hat{x}_i - x_i\right) \quad 3.5$$

The variability associated with the confluency estimates was given by the precision, which was computed from the bias and the RMSE:

$$p_{\text{intrinsic}} = \sqrt{RMSE^2 - bias^2} \quad 3.6$$

The precision of the algorithm for image confluency determination was termed intrinsic precision, as it is a fixed source of variability that couldn't be minimised without further parameter tuning or modification of the algorithm. This was in contrast with the variability introduced by sampling, which was dependent on the sample size (as discussed in the next section).

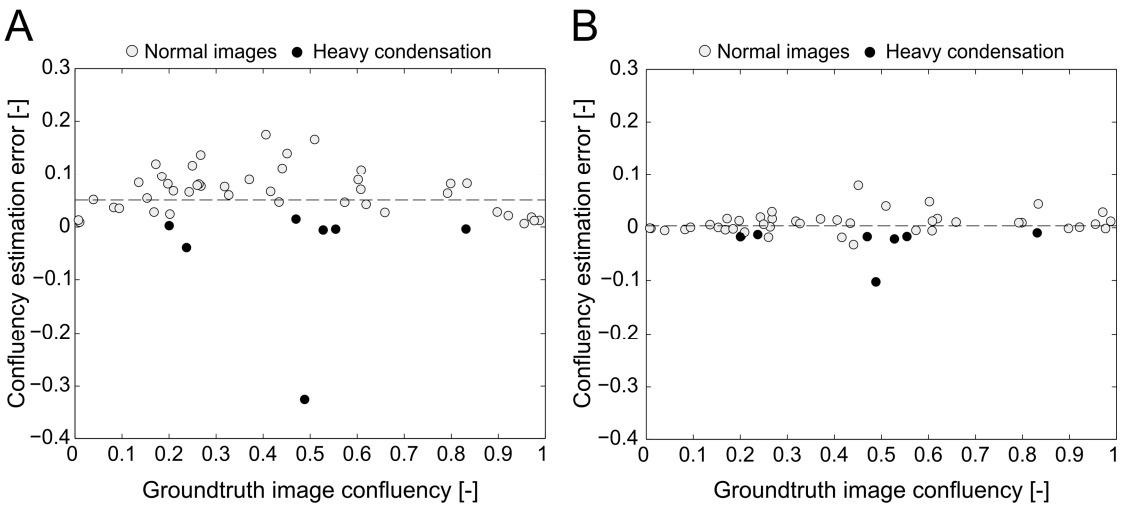


Figure 3-2 Image confluency determination performance. Absolute confluency estimation error for 50 mESC PCM images as a function of the confluency for (A) images processed with the nseSTDEV filter and (B) with the nseSTDEV filter and *post-hoc* halo correction. Light-filled markers represent normal images while black-filled markers denote images with heavy condensation. The dashed line is the mean computed for all 50 images.

The impact of segmentation quality on image confluency estimation was investigated by comparing results obtained based on the segmentation output with and without halo correction (Figure 3-2 and Table 3-1). Without halo correction, confluency tended to be over-estimated in many cases, as shown by a 5.1% mean signed error (bias). In comparison, the additional halo correction step dramatically reduced the bias to just 0.4%. Other metrics followed a similar trend and generally showed that performance was noticeably improved with halo correction.

Table 3-1 Confluency estimation performance based on segmentation output from a normalised soft-edged local standard deviation algorithm with and without halo correction

Metric	nseSTDEV	nseSTDEV + halo correction
Bias [%]	5.1	0.4
RMSE [%]	8.7	2.5
Precision [%]	7.0	2.5
95% confidence interval [%]	[3.1 ; 7.1]	[-0.3 ; 1]

The quality of the estimates did not appear to be dependent on the confluency of the image; even so, extreme ends of the range considered (very low or high confluencies) were less likely to suffer from estimation error as their segmentation was less challenging (small round cells or very large monolithic colonies) compared with images of intermediate confluencies (more likely to contain intricate structures). Another interesting point was that images suffering from heavy condensation during acquisition resulted in more pronounced estimation errors, mostly skewed towards negative mean signed errors (i.e. underestimations). As previously mentioned (see section 2.3.4.3), condensation appears on the lid of culture vessels due to temperature difference between the culture medium (which was kept at 37°C in the incubator) and the environment around the microscope (room temperature). This layer of condensation can perturb the image formation process of phase contrast microscopes, resulting in a loss of overall image contrast and sharpness (Figure 3-3). While halo-correction alleviated this effect, it was still present.

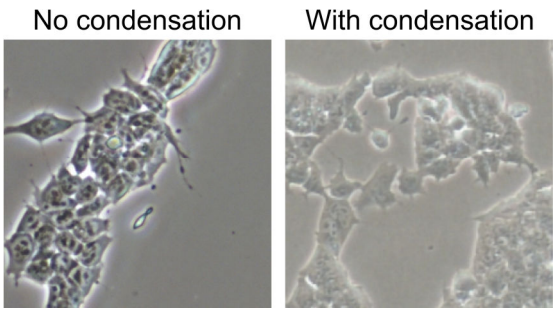


Figure 3-3 Example of image quality differences between PCM images acquired with and without condensation present on the culture vessel lid.

These results suggested that the segmentation outputs of the nseSTDEV filter followed by halo correction were suitable for the estimation of the confluency of a single image.

3.1.1.3 Automated culture confluency estimation

The confluency of a culture was computed as the mean of the image confluency determined for a certain number of images. However, adherent cell cultures are rarely uniform across the whole culture area due to a combination of factors, including the design of the culture vessel, non-optimal seeding protocol or inappropriate handling (Dehlinger et al., 2013; Usaj et al., 2011). Inhomogeneity in local cell number and phenotype within a culture is also intrinsic to some processes (e.g. differentiation) or culture systems (e.g. co-cultures). As a result, the notion of a "representing sample" very rarely applies to adherent cultures. This appears to be widely accepted and acknowledged, as researchers tend to survey a culture at multiple locations during visual inspections. For automating confluency determination, additional factors have to be taken into account. Indeed, sample preparation as well as image acquisition and processing can be relatively lengthy procedures when carried out by a computer as opposed to a human operator. Ideally, the whole culture area would be imaged but this is highly impractical due to the very large number of images to be acquired for most commonly used vessels (Table 3-2). The time taken for image acquisition and processing is critical as the exposition of cells to non-optimal conditions (i.e. no temperature and gas atmosphere control) should be minimised.

Table 3-2 Number of images required at a 10× magnification to cover the whole culture area of commonly used culture vessels. All numbers are based on a rectangular field of view of approximately 1.2 mm^2. This is for illustration purpose only as exact growth area is likely to vary for different manufacturers.

Vessel	Approximate growth area [cm^2]	Fields of view to cover entire growth area
96-well plate	0.32	27
48-well plate	0.95	81
24-well plate	1.9	161
12-well plate	3.8	322
6-well plate	9.6	814
T-25 flask	25	2121
T-75 flask	75	6362
T-125 flask	125	10604

There is therefore a trade-off between throughput (i.e. time required for a measurement) and confluency estimation accuracy. This was investigated by relating this trade-off to random sampling where uncertainty decreased with the number of images averaged. In order to do so, 100 random images of mESC cultured in a well of a 6-well plate were acquired. This sample was assumed a reasonable (or for the purpose of this exercise, to be an exact) approximation of the entire culture area. Based on the standard deviation of the confluency estimate for these 100 images (denoted σ_{100}), the length of the 95% confidence interval was calculated for smaller number of images (n):

$$p_{sampling} = 1.96\frac{\sigma_{100}}{\sqrt{n}} \tag{3.7}$$

The overall precision of culture confluency estimation was dictated by the variability introduced by the sampling (which can be minimised by increasing the number of images considered) and the intrinsic precision of the image confluency determination (fixed variability):

$$p_{\text{culture}} = \sqrt{p_{\text{intrinsic}}^2 + p_{\text{sampling}}^2} \tag{3.8}$$

While previous reports suggested using three or fewer images for culture confluency estimation of a 9.6 cm^2 culture area (Ker et al., 2011; Topman et al., 2011), such low image numbers would have resulted in very high sampling variability according to our calculations. Instead, 10 images were considered a reasonable minimum as it marked the point where the variability introduced by random sampling was less than the intrinsic precision of the algorithm (Figure 3-4). In practice, acquisition time for up to 20 images was found to be reasonable (sampling precision of 1.74%), above which the improvement in precision couldn't justify the decrease in throughput due to diminishing returns. The combined precision of culture confluency estimation using 20 images was thus found to be 3%, which was a ~4-fold improvement over the precision of confluency estimation by human operators for our test case (see section 3.1.1.1)

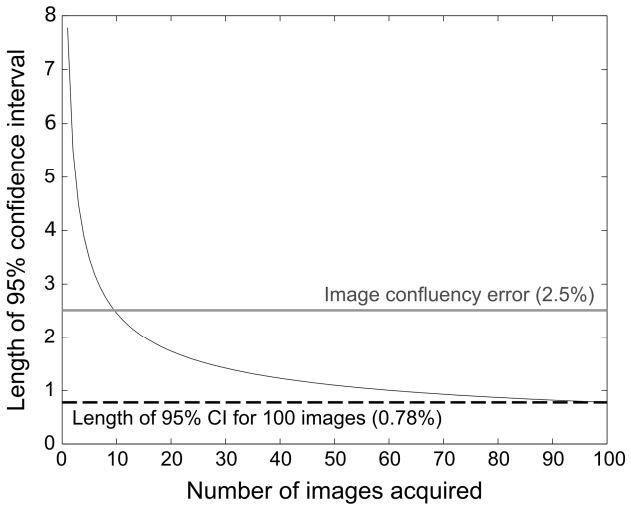


Figure 3-4 Evaluation of the impact of sampling error on culture confluency estimation. The confidence was calculated from the standard deviation of confluency estimated based on 100 fields of view. For comparison, the image confluency error is shown (light grey line).

While this example calculation results would need to be further confirmed by considering other cell lines, culture vessels and imaging protocols, it suggested that the method proposed for automated confluency estimation offered a significant improvement in precision over human researchers while maintaining comparable measurement times.

3.1.2 Morphometric analysis

Cell (and colony) morphology is a key characteristic of an adherent cell culture. Morphological changes were observed in response to flow-induced shear (Sakamoto et al., 2010), thermal shock (Sugimoto et al., 2012) or the perturbation of cells with various chemicals and small molecules (Dong et al., 1998; Jeong et al., 2005; Stroka et al., 2012). When these morphological changes are accompanied by variations in specific cell area, they can be detected by monitoring culture confluency (i.e. measure of the culture area occupied by cells). In other cases, it may be necessary to analyse the shape and properties of individual objects, a process known as morphometric analysis. For this project, two metrics were used to characterise the shape of cellular object: the form factor and the solidity. Both of these can be computed from the output of the segmentation algorithm. In order to do so, properties of individual objects present in the segmented binary mask were calculated using MATLAB's *regionprops* function. Such properties included the area, perimeter, and convex hull of objects (Figure 3-5). The convex hull corresponded to the smallest convex polygon that could contain the cellular object.

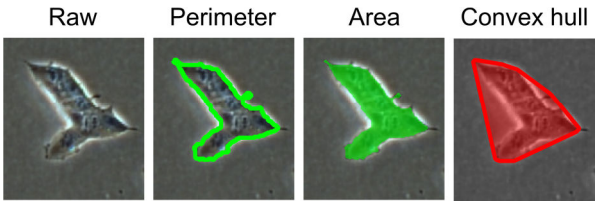


Figure 3-5 Examples of morphological properties that can be extracted from the segmentation output.

The form factor is part of a set of dimensionless numbers that can be used to describe the shape of objects (Soltys et al., 2005). It is computed based on the area and perimeter of objects as follows:

$$FF = \frac{4 \cdot \pi \cdot \text{Area}}{\text{Perimeter}^2} \tag{3.9}$$

The form factor varies from 1 (perfect circle) to about 0.79 (square), and further decreasing towards 0 as perimeter to area ratio increases (e.g. star shape with concave features). Another metric of interest was solidity, a measure of convexity computed as the ratio of the object area to that of its convex hull (Soltys et al., 2005). Values close to 1 therefore denoted objects with high convexity. The solidity decreased as the concavity of the object increased. Both the form factor and solidity were used, in addition to object area, to characterise single cells (i.e. cells shortly after seeding) and colonies.

3.1.3 Augmented fluorescence images

PCM image segmentation can be used to simplify the analysis of fluorescence microscopy images. It is often of interest to spatially relate a fluorescence signal to the location of cellular objects, for example to determine the proportion of cells producing a fluorescent compound. In order to do so, it is necessary to combine the information of two separate images of the same field of view, a process termed registration (Gonzalez and Woods, 2007). Usually, two fluorescence microscopy images are used: one of the molecules of interest and the other of a cell marker (e.g. membrane). Ideally, the latter provides a high contrast image where the boundaries of cellular objects (or that of intracellular compartments) can easily be extracted using simple image processing algorithms, thus providing the information necessary for registration (Ng et al., 2010; Pasquier et al., 2012). However, such an approach can be difficult to apply in practice. While fixed samples usually provide suitable images for the determination of cell boundaries, it is challenging to obtain images of appropriate quality in a live cell imaging settings. In many cases, compounds found in adherent cell culture media, and to a lesser extent the culture vessels themselves, can present varying degrees of autofluorescence (Galdeen and North, 2011). This causes issues as live cell markers usually produce very faint signals (small signal-to-noise ratio), thus precluding the accurate and precise detection of cellular boundaries.

Alternatively, the boundaries of cells can be detected using light microscopy image segmentation techniques such as the one described in this work. By doing so, unnecessary handling of cells is minimised (i.e. addition of cell marker or medium exchange to avoid autofluorescence are both no longer required) and a fluorescence channel is freed, potentially enabling the imaging of additional molecules of interest. It also has the advantage of minimising the exposure of cells to potentially toxic markers and fluorescence illumination. A recent study

showed that using multiple out-of-focus bright field microscopy images could provide a suitable basis for the registration of fluorescence signals (Selinummi et al., 2009). However, this method potentially increase the processing time and requires automated microscopy setups that can precisely control the position of the sample in the z-direction.

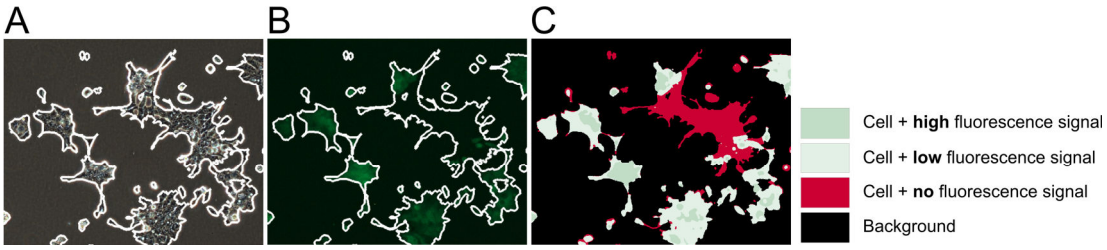


Figure 3-6 Augmented fluorescence image (AFI) constructed by combining PCM image segmentation and fluorescence microscopy. (A) PCM image with the segmentation outcome overlaid in white. (B) Fluorescence image of the same field of view with the PCM segmentation outcome overlaid in white. (C) AFI generated from (A) and (B). Green regions correspond to cell regions that are also associated with fluorescence (dark and light for strong and weak fluorescence signal, respectively). Regions in red correspond to cell region with no fluorescence signal. Black regions correspond to background (non-cellular regions).

Here, an alternative approach is introduced. Fluorescence microscopes are usually also equipped for the acquisition of PCM images. Given a PCM and fluorescence image of the same field of view, the location of the cellular objects can be extracted from the former and the intensity of the fluorescence signal at each pixel from the latter. Indeed, the segmentation mask obtained from the PCM image (Figure 3-6.A) can be used to locate cells on a fluorescence image of the same field of view (Figure 3-6.B), enabling the discrimination between cell-free background and cells with no fluorescence signal, or the classification of fluorescent cells according to their fluorescence intensity (Figure 3-6.C). In this work, this combined representation was termed augmented fluorescence images (AFIs). AFIs facilitate the interpretation and quantification of fluorescence data without the need of whole-cell or membrane fluorescence markers.

3.2 Application examples

Having established the analytical methods in the previous section, they will here be applied to various scenarios relevant to biological fields where adherent cells are commonly used, including biotechnology, drug discovery, and toxicology testing. For each scenario, the experimental setup is described, followed by a discussion of the results obtained, and finally a summary that also outlines possible improvements is given.

3.2.1 Impact of culture medium exchange schedule on mESC growth kinetics

Culture medium exchange is fundamental to adherent cell culture (see section 1.1.3). It enables the replenishment of essential nutrients and small molecules, and the removal of potentially toxic metabolites produced by cells. The exchange rate will be mostly dictated by practical (e.g. number of operators, working hours) and financial (i.e. cost of the culture medium) considerations. Therefore exchanges are often based on a pre-determined schedule. Said

schedules are difficult to determine due to the destructive nature of conventional end-point analytical methods. Unless labour-intensive and expensive serial-sacrificial experiments (i.e. multiple cultures are seeded simultaneously and periodically sacrificed for analysis) are conducted on a regular basis, growth kinetics are at best approximated, at worst unknown. This prevents the establishment of robust, well-defined, and optimal culture protocols.

In this experiment, we investigate the use of PCM segmentation and confluency determination for growth monitoring of mESC cultured under various medium exchange schedules.

3.2.1.1 Experimental setup

Three different medium exchange schedules were investigated: (1) no exchange, (2) one exchange after 48h, and (3) two exchanges, after 48h and 70h. At the beginning of the experiment, mouse embryonic stem cells (mESC e14tg2a, passage < 50) were seeded in 6-well plates at a density of 5×10^4 cells cm^{-2} in 2 mL of culture medium (see section 2.2.1.1 for formulation). Three wells were seeded per condition on separate 6-well plates. When imaging, 20 random PCM images were acquired for each well. The imaging time for a given well plate, including the transport to and from the microscope, was kept to less than 10 minutes. No imaging was carried out during the first 20 hours, in order to allow for the cells to settle down and attach to the growth substrate. In all, cells were monitored for ~90 hours.

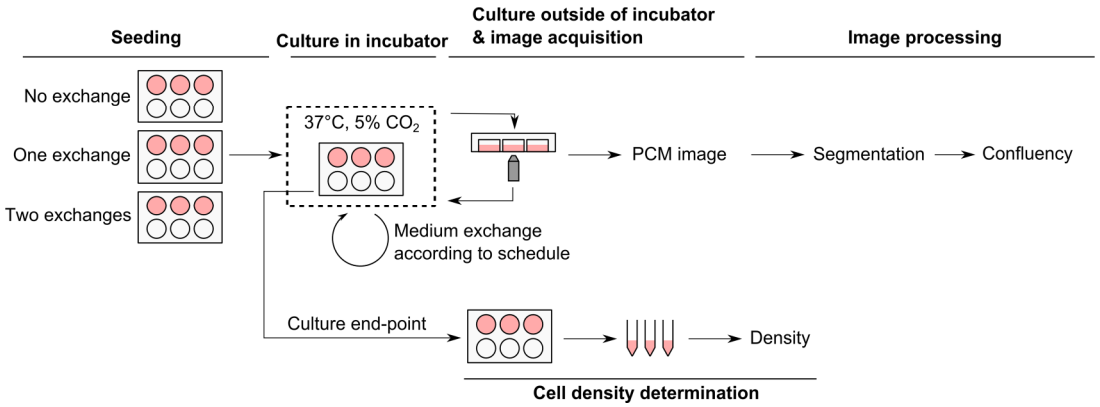


Figure 3-7 Experimental setup for the monitoring of the impact of medium exchange on the growth of mESCs.

Image processing was carried out using an nseSTDEV contrast filter followed by *post-hoc* halo correction with optimal parameter values as reported in section 2.3.2. Culture confluency was determined by averaging the image confluency obtained for 20 random PCM images at a 10× magnification (~1.2 mm^2 field of view). End-point cell density measurements were carried out using an automated cell counter (ViCell, Beckman Coulter, High Wycombe, UK).

3.2.1.2 Results and discussion

Confluency of mESC cultures in 6-well plates was regularly monitored over ~90 hours (Figure 3-8.A). Up to the first medium exchange, the confluency measurements for the three sets of

cultures were very similar (one-way ANOVA, p-value > 0.96). Confluency increased at a slower rate past the 40 hours time point, which might have been an indication that cell growth was impeded by either the depletion of nutrients or the accumulation of toxic metabolites. The rates of confluency increase of the two sets of cultures whose culture medium had been exchanged recovered to levels similar to those observed during the first stage of the culture. This spike resulted in significantly higher confluency measurements (~1.6-fold) after 60 hours when compared to the cultures with no exchange (p-values < 0.001), whose confluency increased at a much lower rate until about 65 hours into the culture. Past this point, the confluency started to decrease, suggesting a loss of viability, with apoptotic and necrotic cells eventually detaching from the growth surface. This was further confirmed by visual inspection of the corresponding PCM images, which showed a large number of free-floating apoptotic bodies (Figure 3-8.B.i).

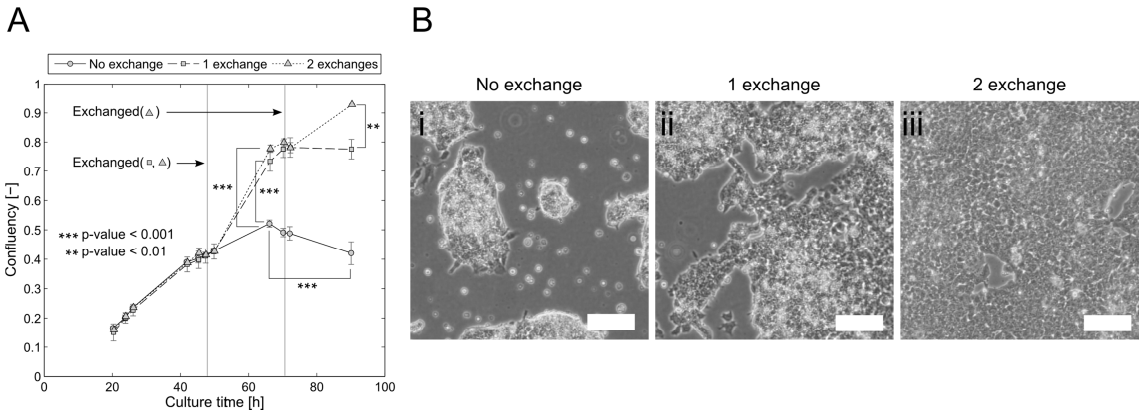


Figure 3-8 Growth kinetics for various medium exchange regimes. (A) Time course study on the effect of medium exchanges on confluency. 20 random PCM images per well (of a 6-well plate), at 10× magnification, were used for confluency determination using the segmentation algorithm (nseSTDEV followed by halo correction). Data shown as mean ± standard deviation (across 3 wells). (B) Representative PCM images of the cultures after 90 hours for the cultures with no exchange (i), 1 exchange (ii) and 2 exchanges (iii). Scale bars are 100 μm.

The two sets of cultures whose medium was exchanged exhibited nearly identical growth profiles until the second exchange. The confluency of the cultures with no further exchange remained constant at a value of ~76% between 70 and 90 hours. The second exchange resulted in another spike in the rate of confluency increase, with an end-point confluency of over 90% at the 90 hours mark, a significant increase (p-value <0.01) over the value measured for the set of cultures with a single medium exchange. This difference in population size between the different conditions tests was also made clear by visual inspection of the corresponding PCM images (Figure 3-8.B.ii and Figure 3-8.B.iii). A one-way ANOVA further confirmed that the number of exchange had a significant effect on end-point confluency (p-value < 10^{-6}). Similarly, end-point cell density determination after detachment of the cells resulted in the same ranking (2 exchanges > 1 exchange > no exchange) and comparable confidence in the effect of the number of exchanges (p-value < 10^{-5}, one-way ANOVA). In this context, the confluency measurements were thus as informative as the cell counts obtained after detachment of the cells.

As the changes in confluency after medium exchanges could have been an artefact of the exchange process (e.g. additional handling, such as aspiration of spent medium and pipetting of fresh medium, resulting in cellular stress and morphological changes), 20 random PCM images per well of the same culture (60 PCM images in total) were acquired before and after medium exchange. The confluency values obtained in both cases were not significantly different (p-value > 0.93, two-sample t-test). This result indicated that the differences observed were most likely imputable to actual increases in culture confluency instead of being an artefact of the medium exchange procedure itself.

3.2.1.3 Summary and outlook

This experiment demonstrated that confluency determination based on PCM image processing was suitable for culture growth monitoring. Measurements were obtained in a matter of seconds and were non-destructive. Because the instruments used for the generation of this data were readily available in cell culture laboratories, this approach enabled the transition from qualitative visual observations to a quantitative and documented framework without incurring significant increase in monetary or time requirements. As such, it can be integrated with existing experimental workflows.

The ability to monitor adherent cell growth non-destructively would enable the thorough characterisation of processes for which information is usually limited to start and end point measurements. Time-course growth data could enable the use of advanced culture strategies based on quantitative measurements of process variables that are usually limited to suspension cultures in instrumented reactors. For example, optimal growth conditions could be devised in order to efficiently expand cells from primary cultures in as few passages as possible to prevent alteration of cell properties (Kretlow et al., 2008; Park et al., 2008; Veraitch et al., 2008).

Continual monitoring also allows for the establishment of experimental protocols in which decision-making is based on quantitative measurements instead of qualitative observations or arbitrary time schedules. Combined with automation systems that are being developed for the manufacturing of large number of adherent cells (Thomas et al., 2008), non-invasive growth monitoring could enable feedback-based control of culture conditions and the application of advanced feeding-strategies that were previously shown to greatly enhance the outcome (e.g. yield, productivity) of processes based on suspension cultures (Csaszar et al., 2012).

3.2.2 Transient cell response induced by environmental shock

The time scale associated with cellular events at a population level (e.g. proliferation or differentiation) is usually of the order of hours or days. However, cell responses of interest might be on a much shorter time-scale than this, and can instead occur in minutes or even

seconds. Moreover, the change of phenotype in response to stimuli can be transient or temporary in that the original phenotype can be recovered later in time (e.g. in the absence of said stimuli). If the experiment is not designed around detecting such short-term cell responses, the latter could be missed due to either the lack of continual monitoring (i.e. analytics limited to end-points measurements) or to the choice of inadequate analytical methods (e.g. cell counts won't inform on transient morphological changes).

In this experiment, culture confluency monitoring was used to detect subtle transient changes in morphology when cells were subjected to stress.

3.2.2.1 Experimental setup

Mouse embryonic stem cells (e14tg2a, passage number < 50) were seeded in 6-well plates at a density of 5×10^4 cells cm^{-2} in 2 mL of culture medium (see section 2.2.1.1 for formulation). Three wells were seeded per condition on separate 6-well plates. Environmental stress was induced by removing the cells from their optimal growth conditions (37° C, 5% CO_2) and leaving them protected from light at room temperature without gas atmosphere control. The durations of environmental shock investigated were 1h, 3h, 6h, and 24h. Optimal growth conditions were restored after perturbation by replacing the cells in a cell culture incubator.

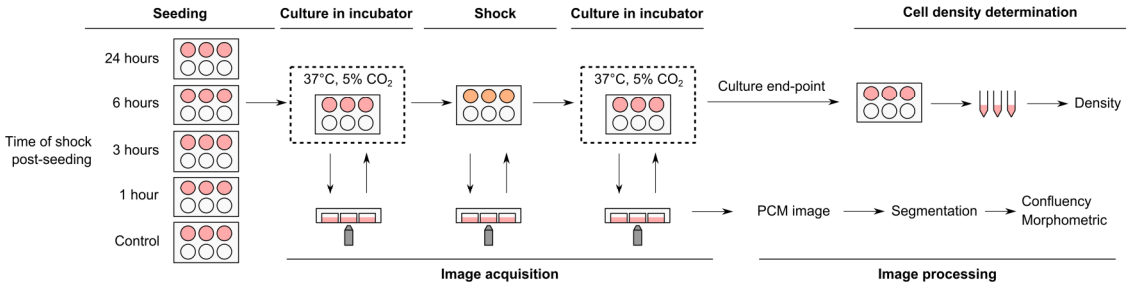


Figure 3-9 Experimental setup for the monitoring of cell response to environmental shock.

Image processing was carried out using the nseSTDEV contrast filter followed by *post-hoc* halo correction with segmentation parameters set to optimal values as discussed in section 2.3.2. Culture confluency was determined by averaging the image confluency obtained for 20 random PCM images at a 10× magnification (~1.2 mm^2 field of view). MATLAB's *regionprops* function was used to determine the mean area of objects on the binary output of the segmentation algorithm. All objects touching the border of the image were discarded prior to analysis.

End-point cell density measurements were carried out using an automated cell counter (ViCell, Beckman Coulter, High Wycombe, UK).

3.2.2.2 Results and discussion

The method chosen to induce cellular stress, whereby cells were left at room temperature without gas control, was first qualitatively evaluated by time-lapse microscopy. When undergoing environmental stress, the cells' visual features visibly changed (Figure 3-10.A). When in optimal growth conditions, the colonies looked smooth and individual cells were difficult to distinguish. In contrast during perturbation, boundaries between cells became much more apparent and colonies' texture appeared rougher. After being placed back into the incubator, the cells recovered the phenotype observed prior to the perturbation, suggesting that the effect, at least in terms of cell morphology and texture, was transient.

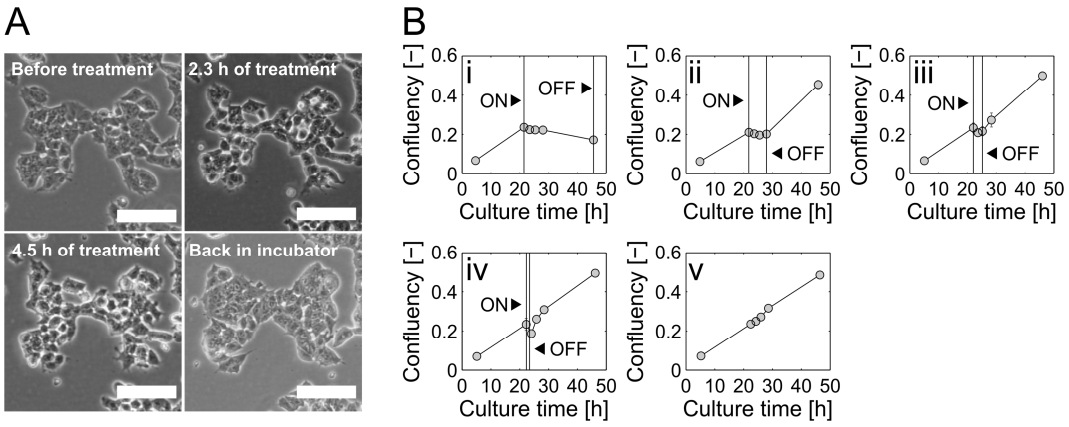


Figure 3-10 Transient cell response monitoring. (A) PCM image of the same colony before, during and after environmental shock. Scale bars are 100 µm. (B) Time course studies of the impact of environmental shock on culture confluency for perturbations of 24 hours (i), 6 hours (ii), 3 hours (iii) and 1 hour (iv) as well as a control case where cells were left in the incubator (v). Vertical lines indicate when the perturbation was started (ON) and when the cells were returned to optimal growth conditions (OFF). 20 random PCM images per well (of a 6-well plate), at 10× magnification, were used for confluency determination using the segmentation algorithm. Data shown as mean ± standard deviation (across 3 wells).

The responses of the cultures to different durations of perturbation were variable. In all cases but the 24 hours perturbation, environmental shock resulted in a significant decrease in confluency within an hour when compared to the control culture (two-sample t-tests, p-values < 0.05). Moreover, confluency remained stable during prolonged perturbation (>1 hour), suggesting that the cells were in growth arrest. 24 hours of perturbation resulted in a decrease in confluency, possibly due to a combination of cellular object shrinking and loss of viability as caused by prolonged unfavourable culture conditions. For perturbation durations of an hour and 3 hours, the measured confluency was no longer significantly different from that of the control cultures as early as 2 hours after restoration of optimal growth conditions (two-sampled t-tests, p-values > 0.1). Similarly, end-point confluency for all conditions but the 24 hours perturbation was not significantly different from that of the control cultures (p-values > 0.9). These results indicated that cells quickly recovered their 'healthy' morphology (at least within two hours after restoration of normal growth conditions) and that short perturbations (< 6 hours) did not adversely affect longer-term cell growth.

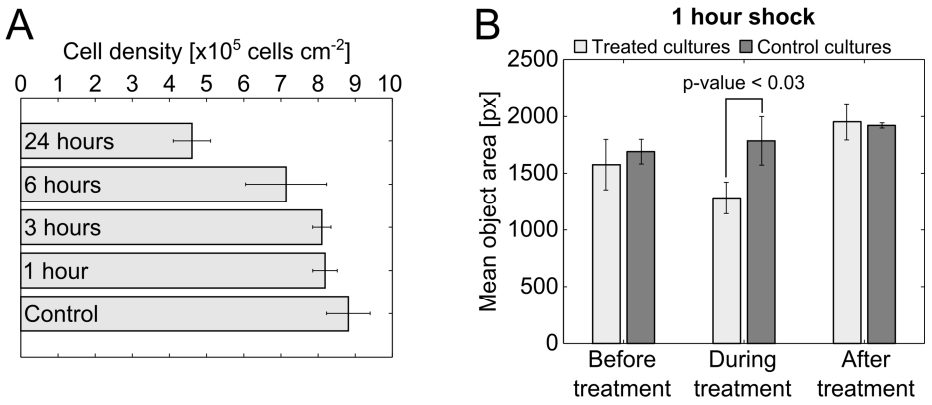


Figure 3-11 Further characterisation of the response of cells to environmental shock. (A) Cell density determined after detachment using an automated cell counter. The data is shown as mean ± standard deviation. (B) Mean object area before, during, and after perturbation for treated and control cultures. The duration of perturbation was of one hour. Area was determined after segmentation and removal of objects touching the border of the image. Data shown as mean ± standard deviation. Three wells and 20 random PCM images per well were used.

The transient nature of the phenotype changes was further demonstrated by determining end-point cell density after detachment of the cells (Figure 3-11.A). With the exception of the cultures that were subjected to the 24 hours perturbation, cell density was found not to be significantly different from that of the control (two-sample t-tests, p-values > 0.14).

In addition to the use of confluency, changes in morphology were also detected by determining the average area of cellular objects (Figure 3-11.B). During perturbation, mean cellular object area was significantly smaller for the treated cultures when compared to that of the control cultures (two-sample t-test, p-value < 0.03). In contrast, the mean area of cellular objects was not significantly different before and after perturbation (two-sample t-tests, p-values > 0.46). These results were thus in good agreement with those obtained with confluency monitoring.

3.2.2.3 Summary and outlook

The results presented in this section highlighted how imaging-based monitoring could help detect and quantify transient cellular responses that would most likely elude conventional end-point measurements. It would be interesting to further investigate the long-term effects of environmental shock on cell phenotype: while the changes in morphology were mostly temporary for perturbations of less than 6 hours, this level of stress is likely to impact other aspects of the cells' metabolism. Indeed, it was previously reported that environmental stress could lead to significant changes in cell properties such as pluripotency or proliferation capability (Veraitch et al., 2008).

The phenotypic changes that were observed in relatively short periods of time out of the incubator might have implications for the use of image cytometers (Kim et al., 2010). These instruments are getting increasingly popular as an alternative to microscopes for high-throughput analysis of live adherent cell cultures. However, their operation is usually relatively slow and they often do not provide any environmental control for the cells. As such, cells might

be left in non-optimal conditions long enough to induce the changes that were quantified during this proof-of-concept experiment, thus potentially introducing biases to measurements made with such instruments.

3.2.3 Cytotoxicity assay for chemical stressor

Adherent cell cultures are often used as *in-vitro* models for cytotoxicology assays whereby cell response to treatment with a potentially toxic agent is measured. Those assays are frequently carried out using fluorescence imaging (Chandler et al., 2011) or impedance-sensing (Xiao and Luong, 2003). In this section, a simple experiment was devised to determine whether PCM image segmentation and confluency determination can be used to measure mESC response to Tunicamycin, a cocktail of antibiotics that is frequently used to mimic the unfolded protein response (also commonly referred to as endoplasmic reticulum stress) associated with various neurodegenerative disease (Tabas and Ron, 2011; Yoshida, 2007) as well as stem cell self-renewal (Blanco-Gelaz et al., 2010; Cho et al., 2009).

3.2.3.1 Experimental setup

This experiment (cell culture and image acquisition) was carried out by visiting student Ana Keser. Mouse embryonic stem cells (e14tg2a, passage number < 50) were seeded in 6-well plates at a density of 5×10^4 cells cm^{-2} in 2 mL of culture medium (see section 2.2.1.1 for formulation). Three wells were seeded per condition on separate 6-well plates. Tunicamycin (Invitrogen, UK) reconstituted in DMSO (Invitrogen, UK) was added to cultures at a concentration of 1 µg mL^{-1}. Control cultures were treated with the equivalent concentration of DMSO (Invitrogen, UK).

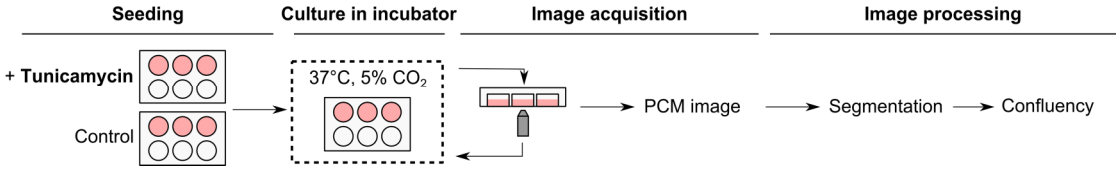


Figure 3-12 Experimental setup for imaging-based cytotoxicity assay.

Image processing was carried out using the nseSTDEV contrast filter followed by *post-hoc* halo correction, with the optimal segmentation parameters discussed in section 2.3.2. Culture confluency was determined by averaging the image confluency obtained for 10 PCM images at a 10× magnification (~1.2 mm^2 field of view). The same 10 fields of view were images for the duration of the experiment (time-lapse imaging).

3.2.3.2 Results and discussion

The confluency of control cultures (treated with DMSO only) and cultures treated with tunicamycin was monitored over 48 hours (Figure 3-13.A). The first significant difference in

confluency between the control and treated cultures was detected 6 hours after addition of the tunicamycin (two-sample t-test, p-value < 0.01). The confluency of the treated cultures then remained stable until at least 25 hours into the culture. In contrast, the confluency of control cultures increased following a profile typically observed for normal mESC proliferation (i.e. approximately linearly with time). These results were confirmed by inspection of the time-lapse PCM images, which clearly showed that the cells in the control cultures proliferated and formed sizable colonies whereas cells in the treated cultures remained sparse (Figure 3-13.B). When the last measurement was taken at 48 hours, the confluency of treated cultures had significantly decreased (two-sample t-test, p-value <0.05) when compared to confluency measurements at the previous time point (~25 hours). This suggested that cell viability decreased, which was further corroborated by inspection of the PCM images, which showed a large number of apoptotic bodies in suspension (Figure 3-13.B).

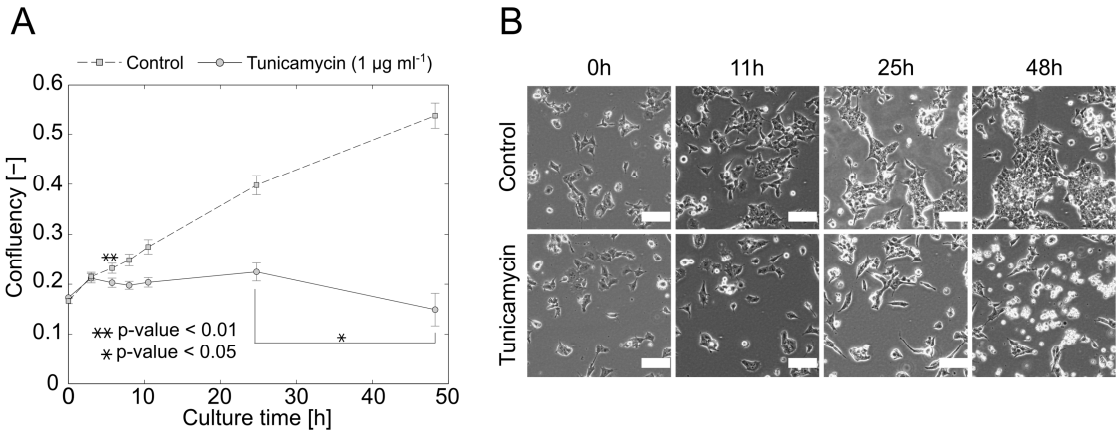


Figure 3-13 Monitoring cell response to a chemical stressor. (A) Time course confluency measurements for cultures treated with 1 µg mL^{-1} of Tunicamycin. Data shown as mean ± standard deviation. Three wells and 20 random PCM images per well were used. (B) PCM images of a representative field of view for control and treated cultures. The same field of view was imaged at different time points. Scale bars are 100 µm.

3.2.3.3 Summary and outlook

The results demonstrated the viability of PCM segmentation and confluency determination for the monitoring of cytotoxic effects on adherent cell cultures. This rather simple experiment showed how the ability to acquire time-course data could help unravel complex cell responses. Indeed, the cell arrest (confluency plateau) followed by apoptosis (confluency decrease) is a response consistent with the current understanding of reticulum endoplasmic stress mechanisms (Yoshida, 2007). This method could be further improved by analysing cell morphology and texture, in addition to culture confluency. Those measurements might be useful for the detection of early cellular stress events that potentially take place much earlier than the 6 hours determined using confluency monitoring, thus providing a refined understanding of this process.

3.2.4 Early and long-term differentiation monitoring

Monitoring of cell differentiation is critical for two reasons. Firstly, it offers a quantitative basis for the understanding of the underlying process kinetics and mechanisms, which in turn can be used for the development and optimisation of differentiation protocols. Secondly, for cells derived from pluripotent stem cells to be used in a clinical context, it is of the highest importance to ensure that the differentiation process proceeds as expected, which requires tools that can be used for quality assurance throughout the entire cell production process.

In this experiment, the segmentation algorithm was used as a basis for the monitoring of differentiation events. Control cultures were grown in expansion medium, which promotes cell proliferation and the maintenance of pluripotency. Two differentiation strategies were considered: spontaneous differentiation by removal of leukemia inhibitory factor (LIF) from the culture medium (a molecule necessary for the maintenance of pluripotency) and directed differentiation towards neuronal lineages using a commercially available differentiation culture medium. Early differentiation events were characterised based on two morphometric attributes (solidity and form factor) as computed from the binary output of the PCM image segmentation algorithm. In addition, a reporter mES cell line that co-expresses GFP with the pluripotency marker *Oct4* was used to detect long-term changes in cell potency. PCM segmentation was used in combination with fluorescence microscopy to monitor spatial and temporal expression patterns.

3.2.4.1 Experimental setup

Mouse embryonic stem cells (Oct4-GiP, passage number < 50) were seeded in 6-well plates at a density of 1×10^4 cells cm^{-2}. Cells were cultured in 2 mL of expansion medium (see section 2.2.1.1 for formulation), spontaneous differentiation medium (same formulation as expansion medium with only 10% FBS and without LIF) and directed neuronal differentiation medium (RHB-A, StemCells Inc., Cambridge, UK). Culture medium was exchanged every 24 hours. Cells were cultured for over 14 days.

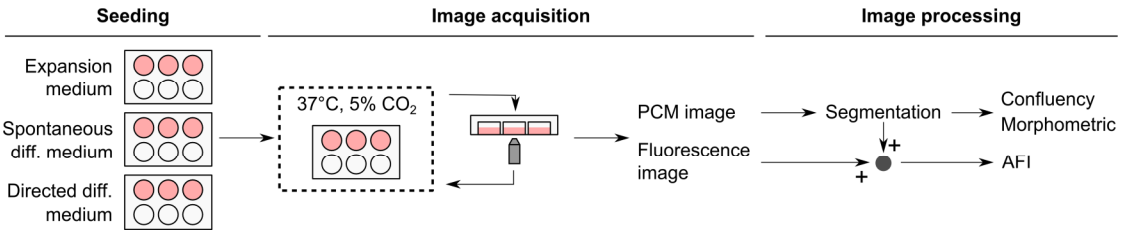


Figure 3-14 Experimental setup for differentiation monitoring experiment. PCM segmentation output is combined with the fluorescence image to generate an augmented fluorescence image (AFI).

The Oct4-GiP cell line expressed GFP under the direction of regulatory elements of the *Oct4* gene, thus allowing to relate GFP expression levels to the potency state of the cells (Ying et al., 2003). For fluorescence imaging, a LED excitation source (CoolLED pE-2, CoolLED, Andover,

UK) was employed to avoid illumination intensity fluctuations usually observed with conventional mercury lamps and thus enabled the comparison of intensity levels between images (Sato and Murthy, 2012). For each field of view considered, a PCM image and a fluorescent image (FITC/GFP) were acquired, both at a 10× magnification. Augmented fluorescence images were generated as described in section 3.1.3. Two measurements were derived from AFIs: the fraction of fluorescent cell pixels (i.e. fraction of pixels detected as 'cell' in PCM images and whose intensity value in the fluorescence image is above a threshold value) and the mean fluorescence intensity of cell pixels. 9 fields of view (three fields of view per well, across three wells) were analysed per condition, for a total of 27 images.

The threshold above which a cell pixel was considered fluorescent (GFP-positive) was set to 0.094 (or equivalent uint8 value of 24, corresponding to ≈ 9% of maximum signal) and that for the separation between low (low-GFP) and high (high-GFP) expressing cell regions to 0.24 (or equivalent uint8 value of 60, ≈23% of the maximum signal). Both thresholds were determined empirically after examination of the fluorescence images and comparison of the signal to background noise. The distinction between low- and high-GFP cell pixels was mostly used for visualisation purposes.

Cellular objects morphology was determined from PCM images. The metrics considered were the solidity (i.e. convexity) and the form factor. These were determined from the binary segmentation output with border-touching objects removed as explained in section 3.1.2.

3.2.4.2 Results and discussion

Changes in morphology associated with early differentiation (~70 hours) were monitored based on the solidity and form factor of cellular objects (Figure 3-15). The mean of both morphometric attributes significantly decreased with time for the control cultures (i.e. culture in expansion medium), from ~0.9 and ~0.75 to ~0.75 and ~0.4 for solidity and form factor, respectively (Figure 3-15.A.ii and iii). This was consistent with visual observations: cells were almost perfectly circular just after seeding (high values for both solidity and form factor) before spreading and eventually forming colonies, at which point objects exhibited complex shapes that tended to be less convex and circular (Figure 3-15.A.i). The morphometric profiles for the cultures in spontaneous differentiation medium showed an initial decrease of the median solidity and form factor, similar to that observed for control cultures, but stabilised earlier and at comparatively higher values of ~0.88 and ~0.63, respectively (Figure 3-15.B.ii and iii). This suggested that cells tended to either remain as individual and round objects (e.g. non-viable cells that can't adhere and spread properly) or grew in well-defined circular colonies (Figure 3-15.B.i). This was in stark contrast with the results for cultures in directed differentiation medium for which both metrics were found to be stable for the period of time considered, at ~0.95 and ~0.75 for solidity and form factor, respectively (Figure 3-15.C.ii and iii). These

findings were in agreement with visual inspections that revealed that cells were not forming colonies but instead remained as small individual objects, sharing similar morphological features with cells shortly after seeding (Figure 3-15.C.i). For all three conditions, time was found to have a significant effect on both solidity and form factor with the exception of form factor measurements during the course of cultures in directed differentiation medium (One-way ANOVA analyses, see Figure 3-15 for p-values).

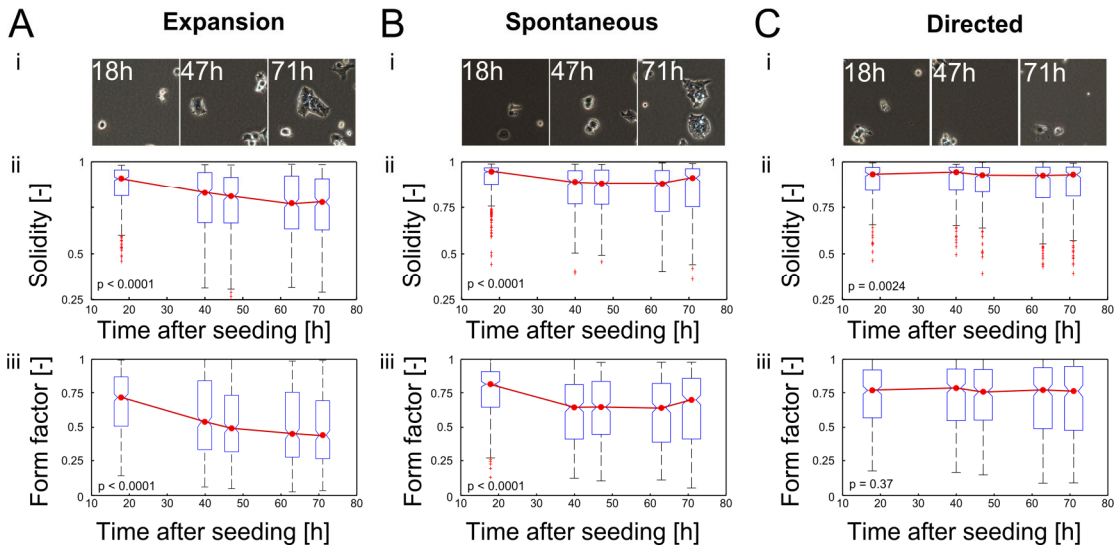


Figure 3-15 Boxplots for the morphometric analysis of early differentiation events. Two metrics were considered: solidity (a measure of convexity) and the form factor. Those were determined on a per-object basis from the binary output of the PCM segmentation algorithm. 3 PCM images per well (9 PCM images in total across 3 wells per condition) were analysed. The p-values shown were determined using a one-way ANOVA with the morphometric attribute as the dependent variable and culture time as the independent variable. The red circles are the median, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The '+' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)]. The top row shows representative PCM images for each condition.

Due to the limitations of the experimental setup used, morphometric analysis could only be carried out for early differentiation events as cells in expansion and spontaneous differentiation media formed large colonies that would fill the entire field of view, thus preventing the detection of their contours and shapes. As an alternative, PCM segmentation was combined with fluorescence microscopy to monitor phenotypic changes associated with long-term differentiation (~14 days). The Oct4-GiP reporter cell line allowed relating GFP expression of the cells with their potency. For each condition, three metrics were considered: fraction of fluorescent cell pixels (Figure 3-16.A), mean intensity of cell pixels (Figure 3-16.B), and mean image confluency (Figure 3-16.C).

At the time of the first measurements (18 hours into the experiment), all three conditions (expansion medium, spontaneous differentiation medium, and directed differentiation medium) only had about 50% of their cell pixels classified as fluorescent (Figure 3-16.A). However, this did not necessarily indicate that only 50% of the cells were pluripotent at the time. Indeed, a study using end-point FACS analysis for fluorescence quantification reported that around 50%

of the cells cultured in optimal conditions for pluripotency maintenance were classified as low-GFP (Veraitch et al., 2008). Given that FACS is generally regarded as much more sensitive than fluorescence microscopy (Jenson et al., 1998) and that the minimum concentration of cytoplasmic GFP required for detection above background levels is relatively high (Niswender et al., 1995), it is likely that the GFP content of the cells was not high enough to be reliably detected by the fluorescence microscopy technique used. This was further corroborated by the rapid increase in the fraction of fluorescent cell pixels as the cultures progressed, most likely due to accumulation of cytoplasmic GFP.

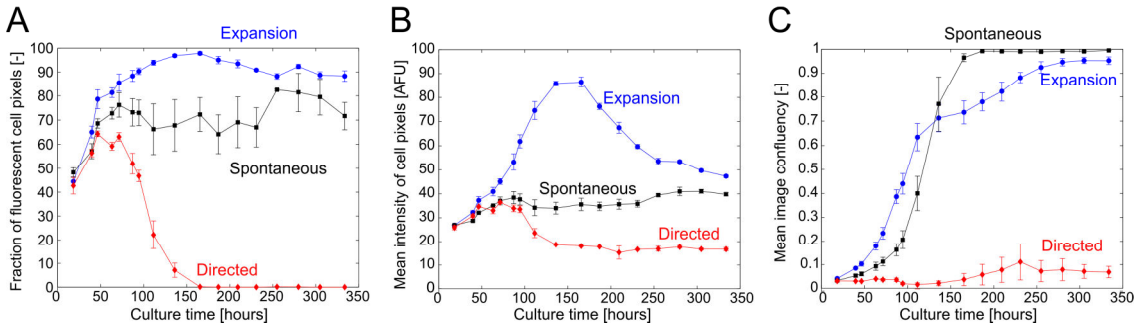


Figure 3-16 Time-course study of mESC differentiation events. Cells were cultured in three culture medium formulations: expansion, spontaneous differentiation, and directed differentiation. (A) Fraction of cells pixels that were also fluorescent. (B) Mean fluorescence intensity of cell pixels. (C) Mean image confluency. In all cases, the data is shown as mean ± standard deviation across three wells. 3 fields of view were imaged per well (9 fields of view per condition across 3 wells).

The cells in expansion medium, which should support the maintenance of pluripotency, reached more than 90% of fluorescent cell pixels within 111 hours of culture. The mean intensity of cell pixels continued to increase past that point, until it reached a plateau at around 90 AFU after 136 hours of culture (Figure 3-16.B). This could potentially be explained as a result of either further accumulation of cytoplasmic GFP or shrinking of cells due to the formation of colonies, reducing the cytoplasmic area. Both phenomena would lead to an increase in the number of GFP molecules per unit area and consequently in the fluorescence signal for the corresponding cell pixels. Even though the fraction of fluorescent cell pixels was stable for the remainder of the culture, a rapid decline in mean fluorescence intensity was measured after 165 hours. This coincided with confluency reaching 70% and a drastic decrease of its rate of increase (Figure 3-16.C). Decreases in both Oct4 expression and proliferative capacity suggested that cells started differentiating despite the presence of LIF in the medium, which should promote pluripotency maintenance (Thomson et al., 1995). Indeed, the proliferation of GFP negative cells as well as the overall decrease in GFP levels was made evident by the inspection of augmented fluorescence images (AFIs, Figure 3-17.A).

Interestingly, high-GFP cells appeared to grow in well-defined, round-shaped colonies until a critical size was reached, after which they remained unchanged for multiple days. It is only when GFP-negative cells started proliferating that the space in-between colonies started filling up. These observations could be related to reports that overgrowth of embryonic stem cells can

trigger differentiation towards neuronal lineages (Reubinoff et al., 2000; Reubinoff et al., 2001). Moreover, the daily medium exchange might not have been frequent enough to maintain adequate levels of LIF (or other essential factors) as the cells proliferated.

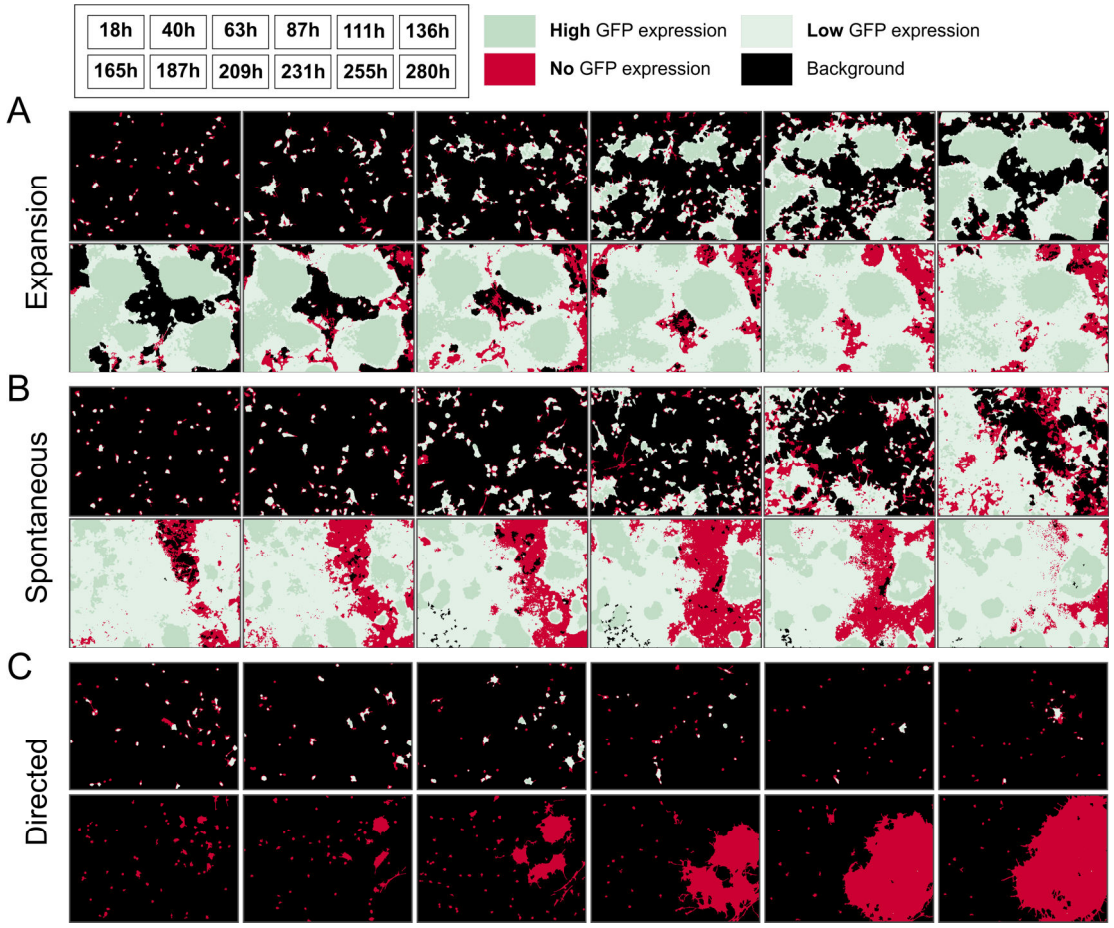


Figure 3-17 Augmented fluorescence images (AFIs) for the visualisation of spatial differentiation patterns. mESC were cultured in three medium formulations: expansion (A), spontaneous differentiation (B), and directed differentiation (C). The time points associated with each image are shown in the top-left insert.

Similarly to control cultures (expansion medium), the fraction of fluorescent cell pixels for cultures in spontaneous differentiation medium also hit a plateau, albeit at a lower values of 70% (Figure 3-16.A). The mean fluorescence intensity of these cell pixels was stable at 35 AFU, which was less than half the maximum value reached for the control cultures (Figure 3-16.B). Interestingly, the growth rate of the cells was initially lower than that of the control cultures (Figure 3-16.C), with cells remaining in small and sparse clumps (Figure 3-17.B). At the 100 hours mark, however, the growth rate soared with the confluency rapidly overtaking that of the control cultures to actually reach full confluency after only 187 hours of culture. Unlike the colonies in control cultures that grew in well-defined colonies until a critical size was reached, cells in spontaneous differentiation medium formed low-GFP clumps without following any obvious pattern. It appeared that the sudden increase in confluency was mostly due to GFP-negative cells rapidly proliferating. The fact that both the fraction of fluorescent cell pixels and mean fluorescence intensity were constant for most of the 14 days suggested that the

cells were in an intermediate state between the potency of the control cultures and the lineage commitment of differentiating cells.

The profiles for all metrics considered were significantly different for cultures in directed differentiation medium when compared to the other two conditions. After only 87h of culture, the fraction of GFP-positive cell pixels plummeted and reached 0% within 165h (Figure 3-16.A). During the same period, the mean fluorescence intensity dropped to levels indistinguishable from background noise (Figure 3-16.B). No proliferation was observed until 187h into the culture (Figure 3-16.C), at which point cells grew into massive GFP-negative colonies that exhibited clear neuronal features (Figure 3-17.C and Figure 3-18).

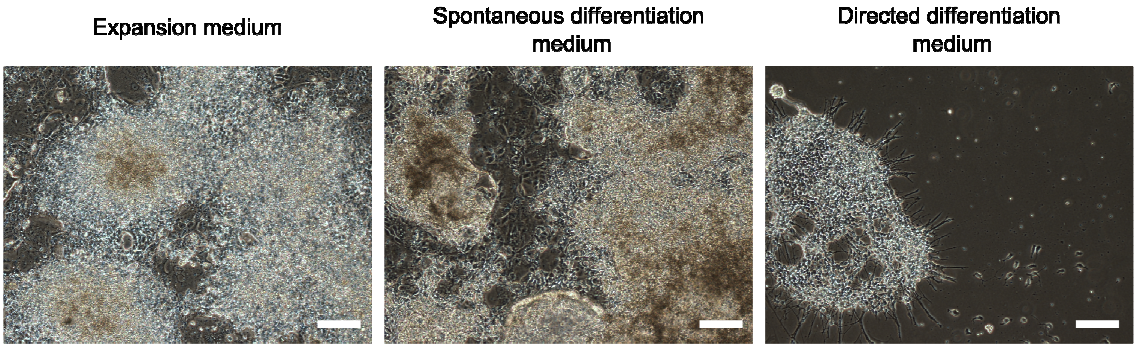


Figure 3-18 PCM images of Oct-4 GiP cells after 280h of culture in expansion medium, spontaneous differentiation medium, and directed differentiation medium. Scale bars represent 150 μm.

3.2.4.3 Summary and outlook

Early differentiation events were monitored using morphological attributes. The results were interesting but highlighted the limitations of the method employed: as soon as objects reached sizes comparable to that of the field of view, morphometric analysis was no longer possible. There would be an advantage in enlarging the fields of view considered by stitching together multiple images of the same region or using a smaller magnification at the cost of reduced resolution. Ideally, the whole culture area should be imaged at each time point, which would require the use of vessels with small culture areas (e.g. 96-well plates) so that acquisition time can remain reasonable. Devices such as the microfabricated bioreactors discussed in chapter 5 would also be suitable.

The combination of PCM image segmentation and fluorescence microscopy data proved successful in monitoring long-term differentiation events. Using a simple experimental setup, it was possible to gain an insight into the changes in cell potency over 14 days of culture. Cells in expansion medium were shown to grow in highly pluripotent, well-defined colonies. There also appeared to be a maximum colony size for pluripotency maintenance above which cells of lower potency located at the periphery of colonies would start proliferating rapidly. In contrast, cells in spontaneous differentiation medium grew in an unorganised fashion. Cells in directed

differentiation medium rapidly lost their potency and only proliferated into large neuronal-like structures towards the end of the 14 days period.

The method suffered from the lack of sensitivity of fluorescence microscopy. As a consequence, early measurements might not be accurate due to the intracellular GFP concentration being too low to be picked up by the camera used. Moreover, GFP is most suitable as a gene expression reporter for cases when a gene is switched on or has its expression enhanced. Due to its long half-life in cells, GFP might not be optimal for the detection of a decrease in expression (Corish and Tyler-Smith, 1999). Instead, the decrease in potency might be better monitored using a marker that gets up regulated during differentiation. GFP was also shown to be cytotoxic in some cases, which might introduce a measurement bias (Liu et al., 1999; Mak et al., 2007).

Despite those limitations, the data clearly showed that this approach has the potential to facilitate the quantification and interpretation of complex fluorescence data. A simple analysis led to the observations of patterns that might be relevant for the understanding of the differentiation process. The quality of the information gained without disruption of the culture could eventually rival that of data obtained using destructive methods such as flow cytometry.

3.3 Towards a method for estimation of cell density from PCM images

In the previous sections, it was shown how useful culture information for a wide range of relevant experimental scenarios could be derived from the results of PCM image segmentation. However, it did not include cell density, which remains the gold standard for cell culture characterisation. Indeed, cell density provides the necessary information to enable calculation of cell metabolic rates for a thorough understanding of cell-based processes (Abaci et al., 2010; Cochran et al., 2006). However, the determination of cell density is mostly limited to end-point measurements, as it usually requires prior detachment of the cells. Ideally, cells would be enumerated directly from PCM images, thus allowing a non-invasive determination of cell density.

It was shown that the number of adherent cells in a culture can be estimated by processing stacks of de-focused microscopy images (Dehlinger et al., 2013). Although the reported performance was reasonable ($R^2 \sim 0.8$ for the linear regression between estimated cell number and actual cell counts), all the results shown were based on non-colony forming cell lines whose cells grow in separate, well-delineated entities, thus facilitating their enumeration. In addition, the need to acquire multiple z-stacks for each field of view significantly increases imaging time and requires a degree of automation that might not be available in a cell culture laboratory. Similarly, a few studies showed that the cell density could be estimated directly from the culture confluency (Juneau et al., 2013; Mölder et al., 2008; Topman et al., 2011). However, the

reported correlation between cell density and confluency was rather weak for cell lines whose area vary greatly during the course of a culture (e.g. colony forming cell lines).

The relationship between cell counts and confluency was investigated by using both PCM and fluorescence microscopy. Multiple 6 well-plates were seeded with mESCs. Before imaging, cells were stained with Hoechst 33342, a live nuclear marker. PCM and fluorescence images of the same fields of view were then acquired (Figure 3-19.A). Because of the potential cellular toxicity of Hoechst, this was done using a sacrificial scheme where cultures were discarded after imaging so that all images were acquired from cultures that have not been exposed to the dye. For each field of view, the PCM image was used for the determination of image confluency using the segmentation algorithm (Figure 3-19.A.i) while the fluorescence image was manually annotated so that nuclei could be counted (Figure 3-19.A.ii). The relationship between the number of nuclei and image confluency was not linear as shown by an adjusted R^2 of 0.86 after linear regression (Figure 3-19.B). Instead, the data was better explained using a quadratic fit.

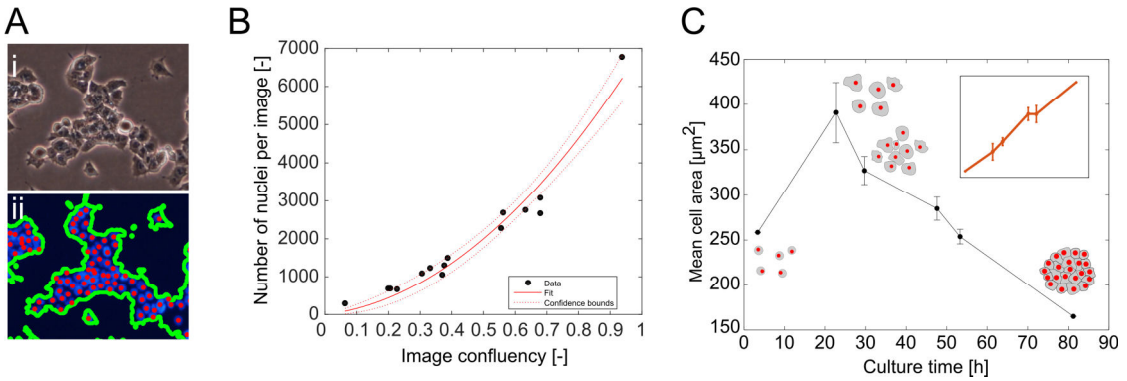


Figure 3-19 Non-linear relationship between image confluency and number of nuclei. (A) Determination of image confluency and nuclei counts using PCM and fluorescence images of the same fields of view. (i) Raw mESC PCM images were processed using the segmentation algorithm to determine image confluency. (ii) Cells stained with Hoechst (a live nuclear marker) were manually annotated (red dots) to determine nuclei count. The green outline represents the segmentation result from (i). (B) Plot showing the relationship between image confluency and the number of nuclei per image. The red line shows a quadratic fit to the data. (C) Mean cell area as the culture progressed. The data is shown as mean ± standard deviation. The cartoon inserts illustrate how cells are spatially organised at different stages of the culture. The plot inserts shows the changes in mean confluency for the period of time considered.

The growth of colony-forming cell lines can usually be broken down in three stages. First, confluency increases while the number of cells remains relatively constant. Then, during an intermediate growth period, the confluency increases linearly with the number of cells. Finally, when approaching high confluency values, small increases in the latter result in a comparatively much larger increase in the number of cells. These three phases are defined by changes in the specific cell area (i.e. the mean area of a cell computed by dividing the total area occupied by cells by the number of nuclei counted) as the culture progresses (Figure 3-19.C). Whereas confluency linearly increased during the course of the culture, changes in specific cell area followed a distinctively different pattern. Shortly after seeding, cells remained small as they had yet to fully spread. Shortly thereafter, an increase in specific cell area was observed, mostly likely due to the cells spreading out, as shown by the relatively constant number of nuclei

counted. As cells formed increasingly tight and dense colonies, the specific area linearly decreased with time, until it reached levels even below the specific cell area observed right after seeding. For this particular system, confluency was therefore a poor predictor for cell density due to the highly variable specific cell area. It is likely that, as reported previously (Topman et al., 2011), the relationship between confluency and cell number for cell lines not forming colonies might have a higher degree of linearity than the one reported here. However, as most non-immortalized cell lines tend to grow in colonies, it might be beneficial to devise a method that enables estimation of cell density from PCM images even for cases with high variations in specific cell area.

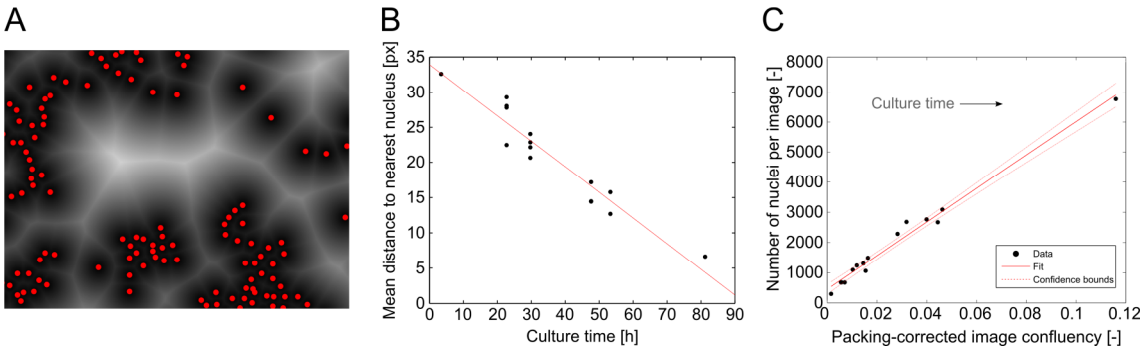


Figure 3-20 Packing-corrected image confluency. (A) Euclidian distance transform of the binary image containing manual nuclei annotations (here shown as red dots). The intensity of a pixel is dictated by the distance to the closest nucleus marker. Small and large distances are represented by black and white pixels, respectively. (B) Mean distance to the nearest nucleus as a function of the culture time. The red line shows the linear regression result, with an adjusted-R^2 of 0.91. (C) The number of nuclei per image as a function of the packing-corrected image confluency (PCC). PCC was computed by dividing the confluency value obtained after segmentation by the mean distance to the nearest nucleus. Adjusted-R^2 of 0.98.

It is clear that a constant correction factor would not be enough to adjust confluency based on variations in specific cell area. Instead, it is necessary to devise a method that would yield a correction factor that depends on the characteristics of the culture. The distance between cells would be one such factor. Based on the manual annotation of the nucleus marker fluorescence image (Figure 3-19.A.ii), this distance can be efficiently estimated using a Euclidean distance transform. First, a binary image was produced where the locations corresponding to annotations was set to 1 and to 0 otherwise. Next, the *bwdist* MATLAB function was used to produce the distance transform: the value of each pixel of the image corresponded to the Euclidean distance from said pixel to the nearest non-zero element or, in this case, the nearest annotated nucleus (Figure 3-20.A). The resulting mean distance to the nearest nucleus was found to decrease linearly during the course of the culture, which indicated that it could be a suitable metric for the correction of confluency to account for changes in specific cell area (Figure 3-20.B). Based on these findings, a new metric termed packing-corrected confluency (PCC) was introduced. PCC accounted for non-linear changes in cell area by dividing the confluency as determined from the output of the segmentation algorithm by the mean distance to the nearest nucleus. The relationship between the PCC and the number of nuclei in an image was found to be linear (adjusted-R^2 of 0.98). On a per-image basis, PCC was therefore found to be a suitable predictor

for cell density. However, this approach required potentially toxic fluorescent markers to be added to the culture so that the distance to the nearest nucleus could be determined, and was therefore invasive. It was thus necessary to devise an alternative method that would not rely on fluorescently marked cells.

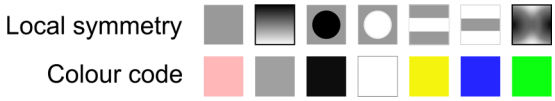


Figure 3-21 The seven Basic Image Features (BIFs) and the corresponding colour code. Additional details about BIFs, and the MATLAB implementation, can be found section 4.2.1.1.

In order to estimate the distance between cells without markers, cellular features had to be detected solely based on the PCM image. This was accomplished by image texture analysis based on the computation of Basic Image Features (BIFs). In short, each pixel of an image was classified as one of seven categories based on local symmetries and structures (Crosier and Griffin, 2010). These categories included dark/bright blobs and lines (Figure 3-21). The computation of BIFs was controlled by two parameters: a scale parameter (σ) that defined the coarseness of the texture analysis, and a threshold parameter (ε), which dictated the fraction of the image that should be considered 'flat' (without texture). For this application, only the 'bright blob' feature was necessary, and was computed using σ=4 and ε=0. Determination of optimal parameter values and feature selection were carried out empirically. More information about BIFs can be found in section 4.2.1.3).

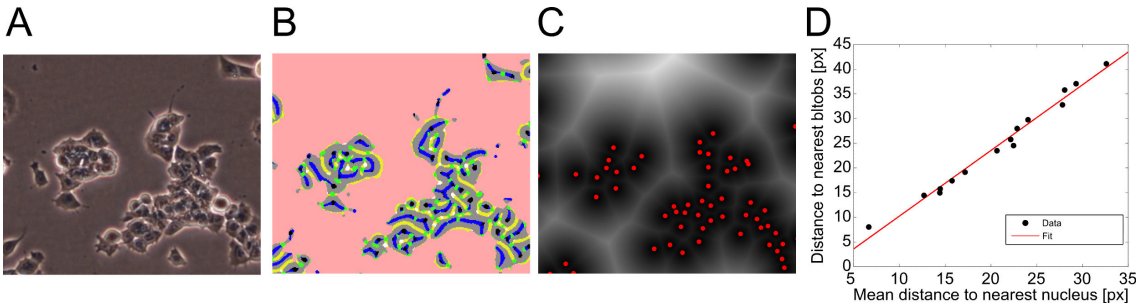


Figure 3-22 Basic Image Features (BIFs) as an alternative to fluorescent markers. (A) Raw PCM image of mESC in a 6-well plate. (B) BIFs computed for the image shown in (A), colour coded as per Figure 3-21. (C) Distance map for the basic image feature corresponding to 'bright blobs', which are shown as red circles. The intensity of a pixel is dictated by the distance to the closest blob feature. Small and large distances are represented by black and white pixel respectively. (D) Distance to nearest blob (as determined as shown in (B) as a function of the distance to nearest nucleus (as determined in Figure 3-20.A). The red line represents a linear regression model. Adjusted-R^2 = 0.99.

Input PCM images (Figure 3-22.A) were first processed using the segmentation algorithm to determine image confluency. The BIFs image was then computed (Figure 3-22.B), and all pixels that were detected as background during segmentation were set to a flat texture. A binary image was generated where all pixels corresponding to the 'bright blob' feature were set to 1, and 0 otherwise. The mean distance to the nearest 'blob' was then computed from the Euclidean distance transform as outlined above for the image annotated using the fluorescence marker (Figure 3-22.C). The mean distance to the nearest 'blob' was found to be linearly correlated with the mean distance to the nearest nucleus (adjusted-R^2 of 0.98), which indicated that the

former could be used as a proxy for the latter, at least for the set of images considered (Figure 3-22.D). Indeed, the PCC calculated based on the distance to the nearest blob had a linear relationship with the number of nuclei in an image that was very similar to that obtained with PCC based on the mean distance to the nearest nucleus (Adjusted-R^2 = 0.99). It was therefore possible to estimate the number of nuclei (i.e. number of cells) in unlabelled PCM images based on PCC.

The next step consisted in investigating whether the ability to estimate the number of cells on individual images would in turn enable the accurate estimation of the culture cell density, which is conventionally determined after detachment of the cells. This implied that the culture should be sampled appropriately during imaging so that the PCC computed would be representative of the entire growth area.

The performance of PCC as a culture cell density predictor was evaluated using a serial-sacrificial methodology. Multiple mESC cultures in 6-well plates were started simultaneously. For each time point, cells were imaged at 20 random locations before being detached and counted using an automated cell counter (ViCell, Beckman Coulter, High Wycombe, UK). As expected from the previous set of results obtained using the fluorescent nucleus marker, the relationship between cell density and culture confluency was found to have a relatively low degree of linearity with an adjusted-R^2 of 0.88 (Figure 3-23.A). When using PCC instead, the relationship was found to be highly linear with an adjusted-R^2 of 0.99. This suggested that PCC was a suitable predictor for cell culture density as determined after detachment.

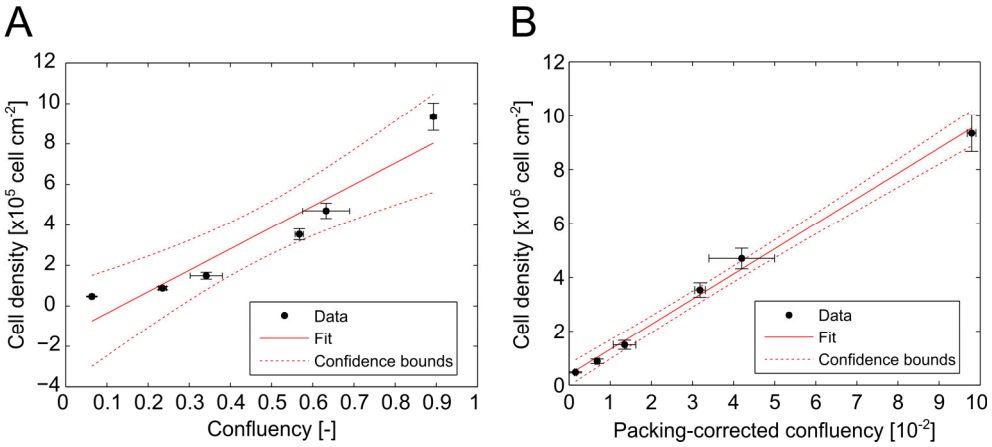


Figure 3-23 Packing-corrected confluency for mESC culture cell density estimation. (A) Relationship between cell density (as determined after detachment of the cells) and confluency. Adjusted-R^2 = 0.88. (B) Relationship between cell density (as determined after detachment of the cells) and packing-corrected confluency. Adjusted-R^2 = 0.99. The data points show the mean while the error bars represent the standard deviation across three wells. For confluency and packing-corrected confluency, 20 random fields of view were averaged for each well.

When used in practice for cell culture monitoring, this method required the creation of a calibration dataset to be used for the determination of the regression coefficients, which in turn enabled the prediction of cell density from PCC values. Such data would be generated by following a serial-sacrificial methodology, similar to the one outlined above. This process was

simulated and validated by using three cultures of mECs in 6-well plates. One culture was used for the generation of calibration data and the resulting regression coefficients were assessed using the testing data computed from the remaining two cultures. This was repeated three times, so that each culture would be used for calibration purposes once. For each iteration, the quality of the estimate was assessed using the mean relative error, the adjusted-R^2 of the regression model between the PCC and cell density, as well as the normalised root-mean-square error (nRMSE), which was computed as following:

$$nRMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{x}_i - x_i\right)}{\sum_{i=1}^{n} x_i}} \qquad 3.10$$

When using PCC instead of confluency for the estimation of cell density, the average mean relative error across the three iterations decreased approximately by a factor 4 (Table 3-3). Similarly, the maximum relative error and nRMSE were improved by more than 7-fold and 2-fold, respectively. Interestingly, the nRMSE obtained for the estimation using PCC, 10.2%, was very similar to previously reported values obtained using confluency for the estimation of cell numbers of non-colony forming cell lines (NIH3T3 and C2C12 with 9% and 10% respectively) (Topman et al., 2011). Thus, the use of PCC made it possible to mitigate the effect of the varying specific cell area characteristic of colony-forming cell lines such as mESCs. Some of the discrepancies between PCC-based estimation and the 'true' cell density values might be attributed to the end-point cell counting method used as a reference. Indeed, multiple factors can impact the variability of cell counts, including inadequate cell handling during and after detachment as well as the intrinsic measurement errors of automated cell counters (Lew et al., 2012).

Table 3-3 Validation of the calibration process for culture cell density estimation based on packing-corrected confluency (PCC).

Method	Iteration	Mean relative error [%]	Maximum relative error [%]	Adjusted-R^2	nRMSE
Confluency	1	-27.3	-289.9	0.91	19.7
	2	-23.5	-252.2	0.89	21.8
	3	-26.4	-251.0	0.88	22.9
	Mean	**-25.8**	**-264.3**	**0.89**	**21.47**
PCC	1	-0.99	26.3	0.98	12.5%
	2	6.2	30	0.98	9.7%
	3	11.5	48.4	0.99	9.5%
	Mean	**5.6**	**34.9**	**0.98**	**10.2%**

These results demonstrated that not only PCC could be used to estimate the number of cells on a per-image basis but that it was also suitable for the estimation of the cell density of mESC

cultures in 6-well plates, based on 20 random fields of view. However, the need to generate calibration data could impede the adoption of such a method. It would be necessary to determine how sensitive said calibration would be to changes in culture conditions (e.g. illumination, culture vessel). While having to establish calibration data for new cell lines is reasonable and expected, the calibration data should be robust and produce suitable estimations for a wide range of conditions. If new calibration data was required for every change in the system, the method would prove unsuitable for monitoring applications.

3.4 Conclusion

The investigations presented in this chapter demonstrated how imaging-based monitoring of adherent cell culture could be applied to relevant experimental scenarios. Using instrumentation that is already in place in most cell culture laboratories, and without incurring additional effort and cost, it was possible to characterise the kinetics of processes that are usually only probed at the end of an experiment. The ability of the presented algorithm to generate such quantitative information is testament to its versatility. Indeed, it was successfully employed to monitor proliferation, growth arrest, cell death, and transient morphological changes in a non-invasive manner. When paired with fluorescence microscopy, it also informed on the patterns forming during neuronal differentiation. Preliminary results suggested that cell density of colony-forming cell lines, the gold standard for cell culture characterisation, could be non-invasively determined from PCM images using the newly introduced packing-corrected confluency, which takes into account how the specific area of cells varies as a culture progresses.

Chapter 4

Trainable segmentation of phase contrast microscopy images

The PCM image segmentation method based on contrast thresholding described in chapter 2 was applicable to a wide variety of relevant cell culture experimental scenarios, as demonstrated in chapter 3. Indeed, the method is expected to generalise well to any application where a single cell type has to be detected against a mostly unstructured background (e.g. surface of a commercial culture vessel). Yet, there are situations for which contrast thresholding is not suitable, including images with significant structural background noise, images of cells whose visual features vary significantly during the course of an experiment, and images of co-cultures.

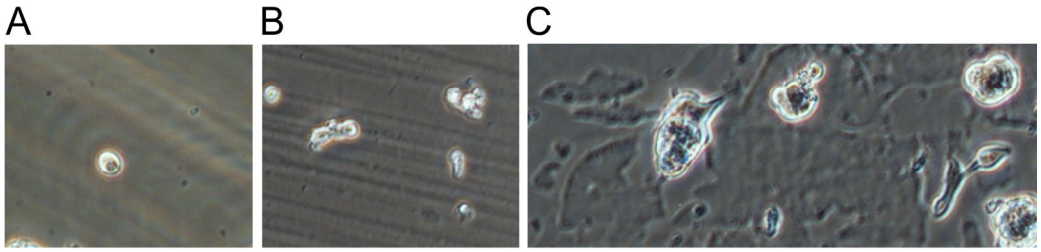


Figure 4-1 Examples of structured background noise. (A) image formation artefact, (B) irregularities (scratches) present on the culture surface or the lid of the culture device, and (C) non-cellular depositions. These images were acquired in the microfabricated culture device described in chapter 5.

Contrast thresholding was found to be tolerant to non-optimal imaging conditions, including uneven illumination across images (section 2.3.4.3). However, severe illumination and image formation artefacts (i.e. artefacts that interfere with the phase contrast image formation mechanisms, such as those inducing diffraction patterns) can exhibit visual signatures that are normally attributed to halo artefacts and cellular objects (Figure 4-1.A). This can lead to the misclassification of these regions as cells. Similarly, contrast thresholding does not always allow discriminating between cells and background artefacts such as scratches (Figure 4-1.B) or non-cellular depositions (Figure 4-1.C). This type of structured background noise is not frequently observed in standard cell culture experiments but might arise when working with small-scale and microfabricated devices (see chapter 5).

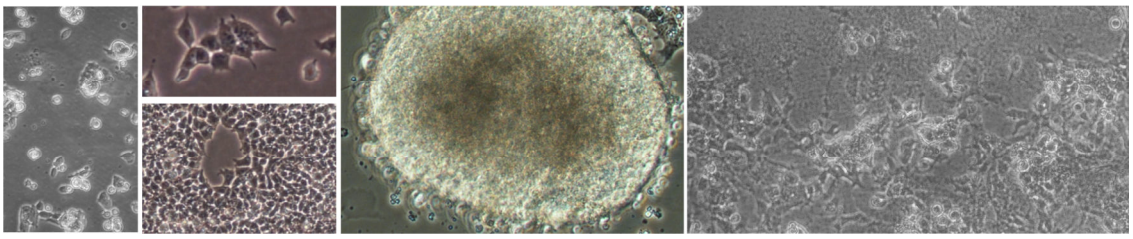


Figure 4-2 A single cell line can exhibit a wide range of visual features. Selection of mouse embryonic stem cells (E14tg2a line) images highlighting the drastic changes in visual features and phenotype that can be observed across applications. Ideally, the segmentation algorithm should be able to accommodate all visual features optimally. The three images on the left were acquired in 6-well plates while the two on the right were acquired in the microfabricated culture device described in chapter 5.

The contrast thresholding algorithm relies on the assumption that background and cells will exhibit a significant difference in intensity homogeneity (see section 2.1.2). This assumption was found to hold true in most cell culture experiments regardless of the cell line used (see section 2.3.4.1). Yet, in specific culture conditions (e.g. differentiation, overgrowth, or flow-induced stress), phenotypic changes might lead to shifts in visual features whereby the interior of cellular objects becomes unexpectedly homogeneous (Figure 4-2). Although this issue can be alleviated by adjusting the value of the intensity threshold parameter (ε, see section 2.1.2.3) accordingly, doing so would likely result in a decrease in performance for 'normal' PCM images. Because the changes in visual features occur gradually during the course of an experiment, it would therefore be necessary to use different variants of the algorithms (and/or of the parameter sets) to optimally segment cells for different stages of said experiment.

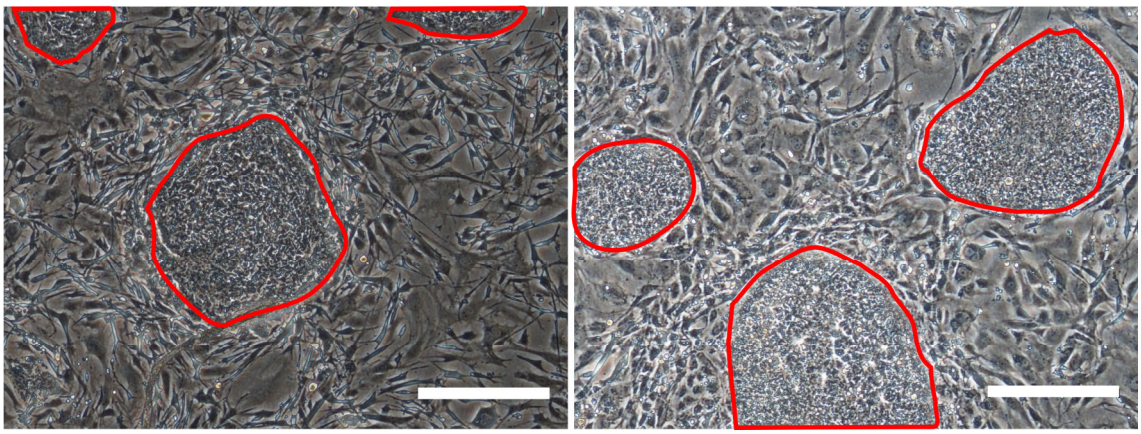


Figure 4-3 PCM images of co-cultures are challenging to segment using conventional approaches. Examples of Human embryonic stem cells (outlined in red) cultured on a background of inactivated mouse embryonic fibroblasts. Scale bars represent 500 µm.

Another situation where contrast thresholding (as well as most other microscopy image segmentation algorithms) would not be suitable is the processing of PCM images of co-cultures. A typical example is the co-culture of Human embryonic stem cells (hESCs) with inactivated mouse embryonic fibroblasts (MEFs). Co-culture is often necessary as MEFs supply factors that promote the maintenance of pluripotency and the proliferation of hESCs (Thomson et al., 1998). Typical PCM images of hESC cultures contain large hESC colonies, which are the objects of interest to be segmented, on a background of MEFs (Figure 4-3). Unlike the PCM images

considered up to this point, there is little differences in intensity homogeneity between foreground (hESC colonies) and background (MEF feeders), thus rendering the contrast thresholding approach unsuitable for the segmentation of hESC co-culture images.

The challenging PCM image segmentation scenarios presented here (structural background noise, unexpected changes in visual features, and co-cultures) highlight the limitations of the contrast thresholding algorithm, but also that of similar image processing approaches in general. Indeed, reliance on both *a-priori* knowledge (e.g. background and foreground homogeneity difference for contrast thresholding) and optimised parameter values make for high performing but specialised algorithms. As shown for the contrast thresholding, it is still possible for such algorithms to achieve good generalisation, as long as the properties and structures of the images to process do not significantly deviate from what is expected based on the *a-priori* knowledge. However, as significant changes in image properties and structures arise, it is necessary to devise new specialised algorithms or, at the very least, to adjust the parameter values accordingly.

Alternatively, a fundamentally different approach to microscopy image processing can be considered, one based on techniques developed within the fields of statistical and machine learning.

4.1 Machine learning methods applied to image processing

Machine learning consist in predicting an outcome, either quantitative or categorical, based on a set of features (Hastie et al., 2009). For this work, only supervised learning methods will be considered, whereby the prediction model (or learner) is built based on training data for which both features and outcome are known. Alternatively, unsupervised learning can be used to build the model based on the features only without knowledge of the corresponding outcomes. Unsupervised learning is thus akin to data clustering. It is important to note that for the purpose of this work, machine learning and statistical learning can be used interchangeably (Breiman, 2001b).

The scope of application for machine learning techniques is very broad (Hastie et al., 2009). A typical example would be the prediction of the price of a property (outcome) based on its location, date of construction, tax code, and internet connectivity (features). This is an example of regression, where the outcome is a continuous variable. This is in contrast with classification where the outcome is categorical (i.e. only takes discrete values). A classic example is the classification of emails as spam or not spam (outcome) based on its word usage statistics and country of origin (features). Individual objects to classify (e.g. an email) are termed ‘examples’.

Machine learning has been extensively applied to image processing and machine vision problems. In fact, applications such as handwritten digits recognition are often used as a benchmark to evaluate new machine learning approaches (LeCun et al., 1989; Lecun et al., 1998; Liu et al., 2003). Handwritten recognition is an example of image classification, whereby pixel intensity values are the features and the digit (0 through 9) is the categorical outcome. In most machine learning approaches, features are encoded in the form of a feature vector of dimensions $1 \times N$, where N is the number of features. In the case of a 16×16 pixels digitized image of a handwritten digit, the resulting feature vector would thus be of dimensions 1×256, or one element per pixel (Figure 4-4). The classifier will then classify the image into one of 10 possible digits based on this 256-dimensional information.

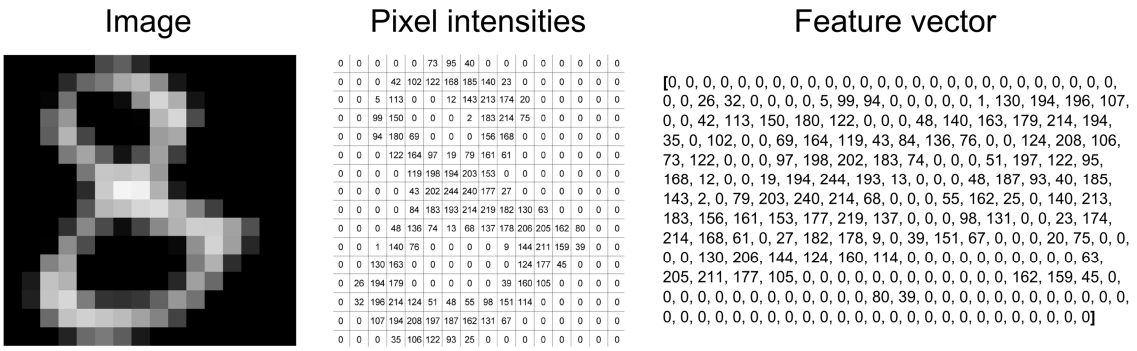


Image
Pixel intensities
Feature vector
[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 26, 32, 0, 0, 0, 0, 5, 99, 94, 0, 0, 0, 0, 0, 1, 130, 194, 196, 107, 0, 0, 42, 113, 150, 180, 122, 0, 0, 0, 48, 140, 163, 179, 214, 194, 35, 0, 102, 0, 0, 69, 164, 119, 43, 84, 136, 76, 0, 0, 124, 208, 106, 73, 122, 0, 0, 0, 97, 198, 202, 183, 74, 0, 0, 0, 51, 197, 122, 95, 168, 12, 0, 0, 19, 194, 244, 193, 13, 0, 0, 0, 48, 187, 93, 40, 185, 143, 2, 0, 79, 203, 240, 214, 68, 0, 0, 0, 55, 162, 25, 0, 140, 213, 183, 156, 161, 153, 177, 219, 137, 0, 0, 0, 98, 131, 0, 0, 23, 174, 214, 168, 61, 0, 27, 182, 178, 9, 0, 39, 151, 67, 0, 0, 0, 20, 75, 0, 0, 0, 0, 130, 206, 144, 124, 160, 114, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 63, 205, 211, 177, 105, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 162, 159, 45, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 80, 39, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0

image (see section 2.1.2.3). Using contrast filter pixel values as features could thus potentially improve the performance of this simple segmentation scheme over one based solely on raw pixel intensity values. The main advantage of pixel-wise segmentation using machine learning, however, is that multiple features can be combined for the classification of a pixel (Figure 4-6). It would therefore be possible to combine both raw pixel intensity and pixel intensity after contrast filtering, which would likely improve the ability of the classifier to discriminate between foreground and background pixels over using either feature individually.

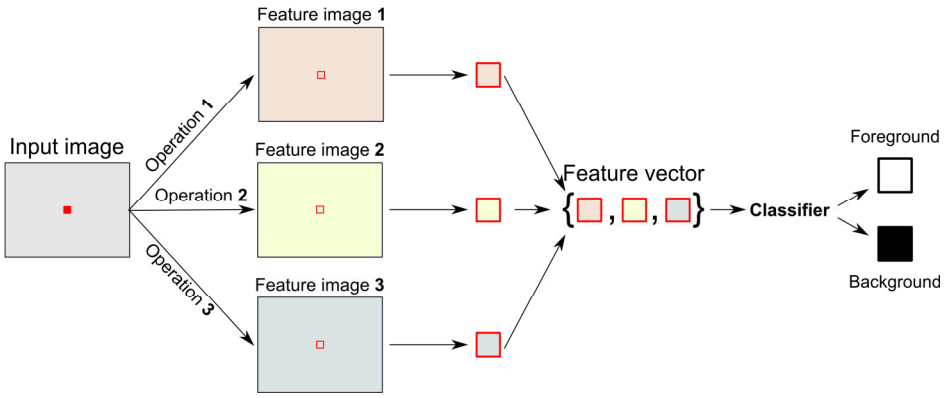


Figure 4-6 Pixel-wise segmentation of an image using multiple features for pixel classification. Each pixel is classified into one of two class labels: foreground or background.

Any type of image feature can in principle be combined in this scheme, including high level descriptors such as local binary patterns (Ojala et al., 1996) or Basic Image Features (Newell and Griffin, 2011). Adding more features can be beneficial but could also lead to overfitting. This is a case of bias and variance trade-off (Hastie et al., 2009), where additional features result in more complex models (lower bias) likely to perform well for the dataset considered but at the cost of lower generalisation to other datasets (high variance). Moreover, increasing the number of features also increase classifier training time as well as memory requirements. These considerations will be further discussed later in this chapter.

Image features can usually be computed at different scales, such as it was the case for contrast filters (see section 2.1.2.3). Varying the scale of a feature often provides additional information about the structure of an image (Figure 4-7). For example, features computed at a small scale can be sensitive to image noise while larger scales reveal larger structures.

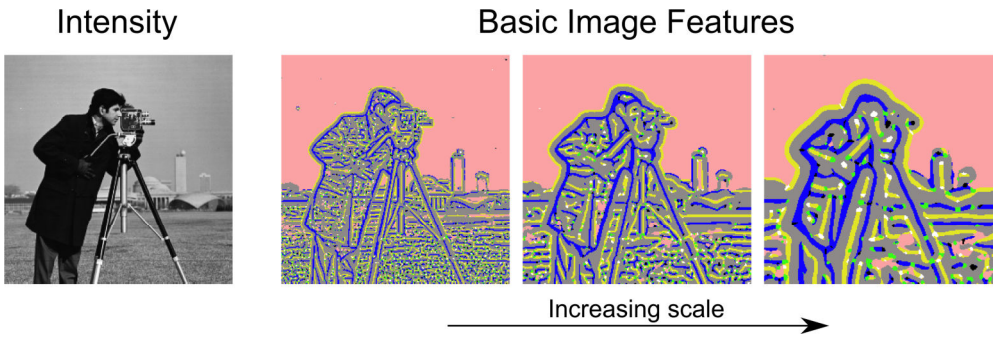


Figure 4-7 Different feature scales reveal different information. Example of Basic Image Features (BIFs) images computed at increasing scale. For details of the colour code and more information about BIFs, refer to section 4.2.1.3.

Exploiting this phenomenon might be helpful when classifying a pixel as background or foreground. Using our previous scheme, the pixel feature vector is constructed by concatenation of the pixel values corresponding to the different combinations of scales and features considered. The resulting pixel feature vector will thus have one element per scale per feature (Figure 4-8). Software packages that support pixel-wise segmentation of images using machine learning techniques, such as Ilastik (Sommer et al., 2011b), are usually based on this scheme or a variation thereof.

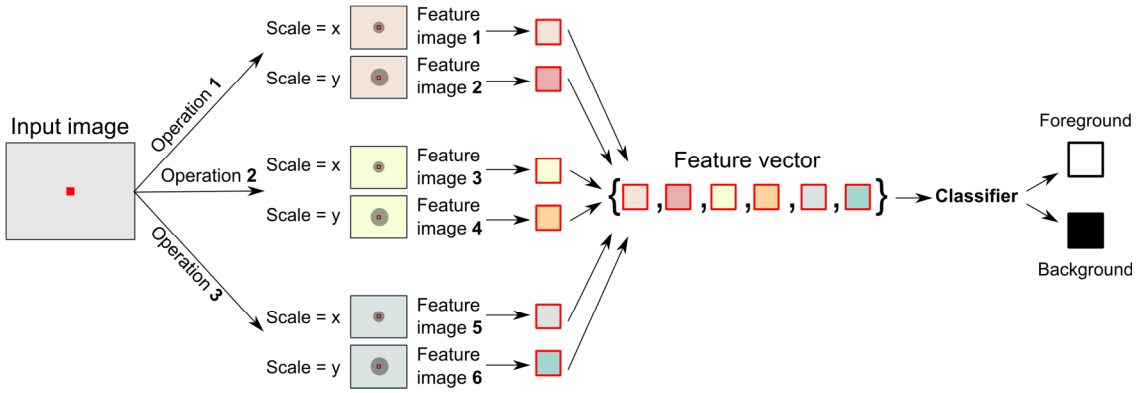


Figure 4-8 Pixel-wise segmentation of an image using multiple features each computed at two scales for pixel classification. Each pixel is classified into one of two class labels: foreground or background.

In this work, we hypothesized that pixel-wise classification could be further improved by taking into account local spatial context and arrangement. The proposed scheme is based on multi-scale image feature local histograms (Figure 4-9). Firstly, different combination of features and scales are computed as above. For each combination, histograms of image features within local windows centred at each pixel position are then computed. For a given pixel, the histograms obtained across each combination of feature and scales are concatenated to produce the pixel feature vector. The discrimination between foreground and background pixels will then be based on the distribution of features in their local neighbourhood. Devising efficient methods for the computation of local image feature histogram is critical to maintain suitable processing times. Various approaches are presented and evaluated later in this chapter (see section 4.2.3).

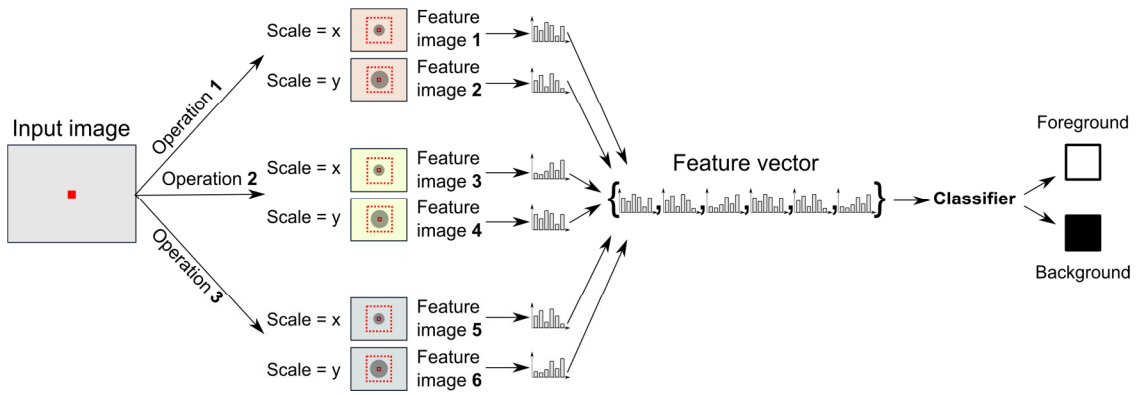


Figure 4-9 Pixel-wise segmentation of an image using multi-scale local image feature histograms. Each pixel is classified into one of two class labels: foreground or background.

4.1.2 Pixel class label prediction using random decision forests

The role of the classifier is to learn the separation of data into classes based on the features provided. There are many criteria to consider when choosing a classifier for a specific application, including the amount of training data, the number of features, the type of features (continuous variables or categorical), the throughput requirements (training and/or prediction time), the nature of the data itself (e.g. noise, correlation), and the degree of generalisation desired (Hastie et al., 2009). The literature on supervised learning is vast and a thorough review is out of the scope of this document. Instead, the main focus will be put on random decision trees classifiers, which will be used for pixel-wise segmentation of PCM images. The reader can find more information about different supervised learning algorithms and methods elsewhere (Bishop, 2006; Hastie et al., 2009).

In principle, any machine learning classifier could be used as part of the segmentation scheme described here and for segmentation applications in general. In practice, however, only a few approaches can be used in the context of pixel classification due to the dimensions of the dataset to classify. Indeed, a typical microscopy image made up of more than a million pixels. Each of these pixels is associated a with feature vector. Supervised classification of image pixels thus requires storing very large data structures in memory and carrying out computationally expensive operations on them for the prediction of the class labels of millions of pixels. As a result, training of the classifier and prediction of class labels for new examples can be very slow. For example, using a support vector machine (SVM, a widely used algorithm for classification) resulted in processing times of up to several minutes per image (Wang et al., 2011). Recently, random decision forests (here abbreviated as random forests, or RF) became a popular choice for pixel-wise segmentation applications. Indeed, RF are memory efficient, fast at both training and prediction phases, have good accuracy, and generalise well (Breiman, 1996; Breiman, 2001a; Criminisi and Shotton, 2013).

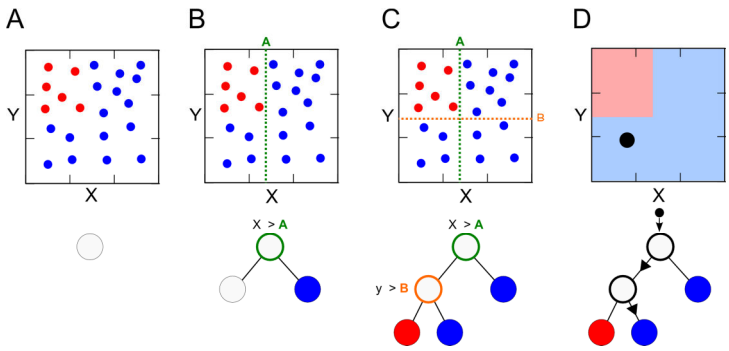


Figure 4-10 Decision tree for classification. Simple example of classification of points into two classes (red and blue) according to their x and y coordinates. (A) The raw data. (B) First best split. (C) Second best split, the data points are now perfectly separated. (D) Once the tree is fully grown during training phase, the class of new examples can be predicted by pushing them down the tree.

RF are part of a family of classification algorithms termed ensemble, whereby prediction is based on the output of a collection of weak learners (Breiman, 2001a; Hastie et al., 2009). In the

case of RF, the weak learners are decision trees (DTs). DTs are a sequence of statements of the form "If $X > A$, then go right, else left", where X is a feature and A is a threshold value learned during training. Each of these statements is a node in the DT, and corresponds to a split of the data (Figure 4-10). For example, one might want to separate two classes of points based on their x and y coordinates (Figure 4-10.A). At this stage, a single node (termed the root node) contains all the data. The goal is to find a single split within this data that would result in two child nodes with maximum separation of the data. The split function used is arbitrary but is often chosen to be axis-aligned, in this case horizontal and vertical lines (Criminisi and Shotton, 2013). The best split is determined by minimisation of a loss function, usually the Gini index (Criminisi and Shotton, 2013; Hastie et al., 2009). The Gini index is minimum for pure nodes (i.e. nodes that contain only examples of a single class) and maximum for nodes that contain equal numbers of examples from each class. Minimising the Gini index thus results in an optimal separation of the data across the child nodes (Figure 4-10.B). This splitting process is then repeated for the child nodes (Figure 4-10.C), growing the tree until a stopping criterion is met, usually a pre-defined tree depth or a threshold on the number of samples arriving at a node after splitting. Terminal nodes (also called leaf nodes) are nodes without children. Once the tree has been fully grown during the training phase, it can be used to predict the class of new samples by pushing the data downwards from the root node (Figure 4-10.D).

Unlike other classifiers that are often described "as black boxes", DTs present the advantage of being easily interpretable. However, DTs are weak classifiers often characterised by poor accuracy, low bias (due to their low complexity), and high variance (Criminisi and Shotton, 2013; Hastie et al., 2009). However, RF combine a large number of DTs with the aim of producing an ensemble model with reduced variance while retaining the low bias of individual DTs. Each tree of the forest is constructed based on subsets randomly sampled with replacement from the original training data, a process termed bootstrap aggregating (or bagging). A given DT is thus likely to contain repeated examples (Figure 4-11.A). The variance is further reduced by randomly sampling the feature candidates for best split selection at each node of the trees (Breiman, 2001a; Hastie et al., 2009). The prediction for a new example is done by pushing its features down all DTs comprising the RF and taking the majority vote across all responses (Figure 4-11.B). Moreover, the classification confidence can be estimated by calculating the proportion of DTs that predicted the majority vote.

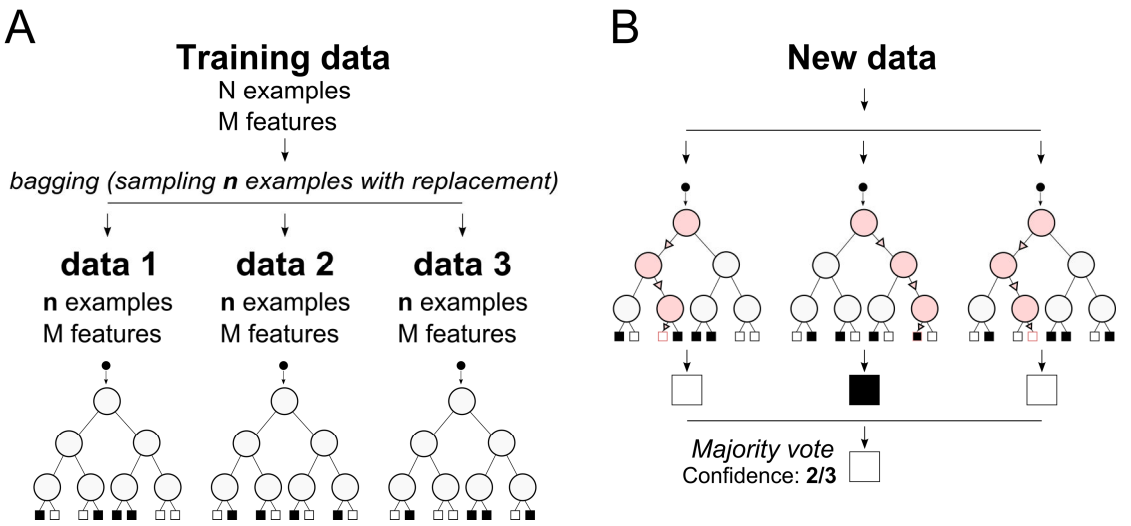


Figure 4-11 Training and prediction phases of a random forest classifier for pixel classification (background vs foreground). (A) The training phase. The training data is sampled with replacement to construct the different trees. (B) The prediction phase. New data is pushed down each tree and the majority vote dictates the final classification outcome. Classification confidence is given by the fraction of the trees that agree with the majority vote. White and black squares represent foreground and background pixels, respectively.

The ability of RF to train and predict quickly for large datasets make them particularly suitable for pixel-wise segmentation applications. Indeed, it was adopted with great success by most software packages supporting trainable segmentation, including Ilastik and the Weka trainable segmentation plugin for Fiji/ImageJ (Hall et al., 2009; Sommer et al., 2011a). The low computational complexity of RF (including memory requirements) compared to other types of classifiers such as support vector machines (SVMs) makes them more appropriate for the pixel-wise segmentation of large images such as those considered in this work. Moreover, in contrast to SVMs, RF has usually less critical parameters to tweak in order to achieve high performance.

It was thus of interest to investigate how trainable segmentation approaches based on a random forest classifier would perform for the task of PCM images segmentation. First, methods for the computation of local image feature histograms are presented and compared (section 4.2.3). Trainable segmentation is then applied to the segmentation of the same mESC PCM images than those used for the validation of the contrast thresholding algorithms in chapter 2 (section 4.3.2). This represents a simple segmentation task (i.e. foreground versus background). Later, similar approaches are applied to a more complex segmentation scenario consisting in PCM images of Human embryonic stem cells co-cultured with mouse embryonic fibroblasts (section 4.3.3). In this case, the task is to discriminate between two very similar foreground objects.

4.2 Computational methods

4.2.1 Features computation

Three types of features were considered: pixel intensity of the raw PCM image, pixel intensity after application of a contrast filter, and basic image features. Each was computed as outlined in

the following sub-sections. In the context of this work, the scale of a feature refers to the granularity at which the feature is computed, with a small scale being sensitive to minute details (and potentially noise) while a large scale might inform on more global image structures (Figure 4-7).

4.2.1.1 Intensity features

Intensity features were computed directly from the raw PCM image. Computation at a given scale was done by blurring the raw image by convolution with a Gaussian kernel of standard deviation equal to the scale.

4.2.1.2 Contrast features

For contrast features, the image was first processed using a normalised soft-edged standard deviation (nseSTDEV) filter as described in section 2.1.2.3. The pixel intensities of the resulting image were used as features. The scale corresponded to the σ parameter of the contrast filter.

4.2.1.3 Basic Image Features

The computation of basic image features (BIFs) consisted in classifying pixels of an image into one of seven categories depending on local structures and symmetries (Griffin et al., 2009). The resulting image whose pixels take one of seven values resemble a primal sketch of the original image. Pixel classification during BIFs computation is based on the responses to a bank of Derivative-of-Gaussian (DtG) filters (Figure 4-12).

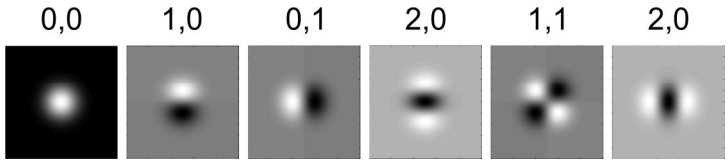


Figure 4-12 Bank of Derivative-of-Gaussian (DtG) filters used for the computation of BIFs. The text above each filters show the order of the filter (*x*,*y*).

The bank of filters included one zeroth-order, two first-order and three second-order DtG filters, each of the same scale (standard deviation) σ. The response of the convolution of the image I with one of the DtG filter was denoted c_{ij} where i and j represented the order in the x and y directions, respectively. For example, c_{20} was the response to the DtG filter with a second order derivative for the x direction and a zeroth order derivative for the y direction. In total, the responses c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, and c_{02} were considered. Convolution was computed efficiently using the separability property of Gaussian kernels (i.e. for each dimension, the convolution was carried out separately using 1-dimension kernels of standard deviation σ, termed $G^i_{x_\sigma}$ and $G^j_{y_\sigma}$ for the x (of order i) and y (of order j) direction kernels, respectively):

$$c_{ij} = \left[I \otimes G^{i}_{x_\sigma} \right] \otimes G^{j}_{y_\sigma} \quad 4.1$$

Scale normalised response $s_{i,j}$ was then computed as follows:

$$s_{ij} = \sigma^{i+j} c_{ij} \quad 4.2$$

The following intermediate calculations were then performed:

$$\lambda = s_{20} + s_{02} \quad 4.3$$

$$\gamma = \sqrt{\left(s_{20} - s_{02}\right)^2 + 4 s_{11}^2} \quad 4.4$$

Both λ and γ were computed for each pixel of the input image I. Pixels were then classified based on which term listed in Table 4-1 resulted in the largest value. BIFs computation was controlled by two parameters: (1) the standard deviation (feature scale) of the DtG filters σ, and (2) a threshold value ε that specified the fraction of an image that should be considered as flat (i.e. no specific structure). For this work, ε was kept constant at 0.03, which was empirically found to produce good results at all the scales considered.

Table 4-1 Classification of BIFs according to which term results in the largest numerical value. The colour codes are used for graphical representation of BIFs image.

Term	Associated label if largest	Structure type	Example	Colour code
$\varepsilon \cdot c_{00}$	1	Flat texture		
$\sqrt{c_{10}^2 + c_{01}^2}$	2	Slope		
λ	3	Dark blob		
$-\lambda$	4	Bright blob		
$\frac{\gamma + \lambda}{\sqrt{2}}$	5	Bright line		
$\frac{\gamma - \lambda}{\sqrt{2}}$	6	Dark line		
γ	7	Saddle point		

4.2.2 Local feature histograms

Two types of local feature histograms were considered (Figure 4-13). Discrete histograms were used to encode the frequencies of discrete labels represented by integers. The bins were centred on the value of the label. A typical example was histograms of Basic Image Features (BIFs), whose values can be integers from 1 to 7. Similarly, it can be used to compute the frequencies

of all possible grayscale levels of an 8-bit image (256 possible integer values). In contrast, interval histograms were used for continuous variables with bins defined as intervals.

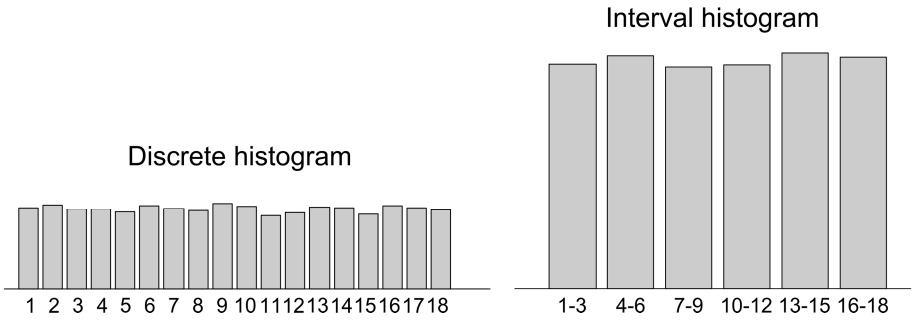


Figure 4-13 The trainable segmentation scheme employed either discrete or interval histograms. 1000 integers ranging from 1 to 18 were randomly generated and binned using discrete 18-bin (left) or interval 6-bin (right) histograms.

Given a feature image (e.g. BIFs or local contrast image), the goal was to generate local histograms of feature values within a given neighbourhood size (*patchSize*) for each pixel of the image. The result of this computation was an $N \times bins$ matrix where N was the number of pixels of the feature image and *bins* the number of bins of the histogram constructed. Unless otherwise specified, the histograms were normalised so that they were independent of the size of the region considered.

Four methods were evaluated for the construction of local feature histograms, as described in the following sub-sections: sliding windows, integral histogram, convolution with a square kernel, and convolution with a (soft-edged) Gaussian kernel.

4.2.2.1 Hard-edged using sliding windows

Hard-edged local feature histograms were constructed using a modified version of MATLAB's built-in *colfilt* function. Firstly, the feature image was padded with zeros, which were ignored during the histograms computation. The feature image was then divided in blocks to ensure memory-efficient processing. The optimal number of blocks was determined using MATLAB's built-in *bestblk* function. For each block, the local neighbourhood of each pixel is stored in a matrix built using the *im2col* function. The dimensions of the neighbourhood were *patchSize* × *patchSize*. The resulting local neighbourhood matrix was thus of dimensions $Nb \times$ (*patchSize* × *patchSize*) where Nb was the number of pixels in the block being processed. The histogram of each neighbourhood was computed using the built-in *histc* function before being stored in the result matrix. The dimensions of said result matrix were $N \times bins$ where N was the number of pixels of the feature image and *bins* was the number of bins of the histogram. The sliding windows approach could be used to construct both interval and discrete histograms.

4.2.2.2 Hard-edged using integral histogram

Integral histogram is an efficient method for the rapid construction of histograms (Porikli, 2005). It is based on a principle similar to the previously described integral image (Viola and Jones, 2002). In short, the histogram was built iteratively pixel by pixel. The simplest approach was termed 'string scan' and consisted in scanning an image from left to right, top to bottom. The integral histogram for a given pixel was computed by incrementing the bin values of the integral histogram determined at the previous step by the amount corresponding to the current pixel. This way, the integral histogram at a given location was a representation of all the histograms between that location and the origin of the image. Alternatively, the integral histogram could be constructed using wave propagation whereby the value at a given location was computed based on three neighbours (left, upper-left and upper neighbours). All the information necessary to compute all possible image histograms was readily encoded in the integral histogram.

The main advantage of this approach was that the histograms were obtained using four lookups into the integral histogram, one for each of the corners of the region, regardless of the size of the region of interest (Porikli, 2005). In this work, the *vl_inthist* (integral histogram computation) and *vl_sampleinthist* (integral histogram lookup) of the VLFeat open-source library (Vedaldi and Fulkerson, 2010) were used. The integral histogram method could only be used to construct discrete histograms.

4.2.2.3 Hard-edged using convolution

Hard-edged local feature histograms with a relatively small number of bins were efficiently computed using convolution. First, a binary mask *b* of dimensions equal to that of the input image was created for each of the histogram bins. When constructing a discrete histogram, elements of the binary mask were set to one if the corresponding pixel on the feature image was equal to the bin value (i.e. its centre), zero otherwise (see equation 4.5). In the case of interval histograms, elements of the binary mask were set to one if the corresponding pixel on the image was within the interval defined by the edges of the bin, zero otherwise.

$$b^{(k)}(x,y) = \begin{cases} 1 & \text{if } I(x,y) = k \\ 0 & \text{otherwise} \end{cases} \text{ for },k = 1,2,\ldots,\text{bins} \qquad 4.5$$

Each of the binary masks was then convolved with a square kernel *g* whose all elements were set to one and of dimensions corresponding to the neighbourhood of interest. The results of the convolution were images *c* (one per histogram bins).

$$c^{(k)}(x,y) = b^{(k)}(x,y) \otimes g(x,y), k = 1,2,...,bins \qquad 4.6$$

The local feature histogram vector H for a given pixel of the feature image was obtained by concatenating the values obtained from the convolutions for that position.

$$H(x,y)=\left[c^{(1)}(x,y),c^{(2)}(x,y),c^{(3)}(x,y),...,c^{(bins)}(x,y)\right] \quad 4.7$$

The convolution approach could be used to construct both interval and discrete histograms.

4.2.2.4 Soft-edged using convolution with a Gaussian kernel

Unlike hard-edged histograms that tally up the frequencies of features in a local neighbourhood uniformly regardless of the distance from the central pixel, soft-edged histograms put more weights in pixels in close proximity of the central pixel. Soft-edged feature histograms were computed by convolution as described for hard-edged histograms (section 4.2.2.3). The only difference being that the kernel g used for convolution was a Gaussian kernel of standard deviation equals to *patchSize* / 2.

4.2.2.5 Statistics of local histograms

An alternative to the use of local histograms bins as features for classification was instead to compute statistics of said local histograms. These statistics were the mean, variance, standard deviation, skewness, and kertosis.

The mean of a histogram was computed as follows:

$$\overline{x}=\sum_{i}^{N} f_i \cdot c_i \quad 4.8$$

Where N was the number of bins, c_i was the centre of the i-th bin and f_i its value. The variance could then be computed as following:

$$\text{var}=\sum_{i}^{N}\left(x_i-\overline{x}\right)^2 \cdot f_i \quad 4.9$$

And the standard deviation of the histogram σ can then be directly calculated as follows:

$$\sigma=\sqrt{\text{var}}=\sqrt{\sum_{i}^{N}\left(x_i-\overline{x}\right)^2 \cdot f_i} \quad 4.10$$

The skewness of the histogram was computed as following:

$$\text{skewness}=\frac{\sum_{i}^{N}\left(x_i-\overline{x}\right)^3 \cdot f_i}{\sigma^3} \quad 4.11$$

Finally, the kurtosis of the histogram was calculated as follows:

$$\text{kurtosis} = \frac{\sum_{i}^{N}\left(x_i - \bar{x}\right)^4 \cdot f_i}{\sigma^4} \qquad 4.12$$

The statistics of histograms feature vector was constructed by concatenation of these five metrics for each pixel of the image considered.

4.2.3 Computation time for local image feature histograms construction

The impact on computation time of various neighbourhood sizes, image sizes, and number of bins was assessed. All tests were performed on a desktop computer equipped with an Intel core i7-2600 CPU at 3.4 Ghz and 16 GB of ram. Computations were not multi-threaded to ensure fair comparisons between the different methods. In each case, the process was repeated five times to minimise the effects of spikes in computation time due to unrelated CPU usage.

4.2.3.1 Impact of neighbourhood size

A 1000×1000 random integer matrix with elements varying from one to the number of bins considered (7, 28 and 56) was generated. The size of the local neighbourhood was varied from one to 9000 pixels. Time to construct one local feature histogram per matrix element (1'000'000 total) was determined using the tic()/toc() functions in MATLAB. Each histogram had 7, 28 or 56 bins.

4.2.3.2 Impact of image size

Random integer square matrices of size 10×10 (100 elements) to 1510×1510 (~2.3×10^6 elements) with elements varying from one to 28 (number of bins) were generated. Time to construct one local feature histogram per matrix element was determined using the tic()/toc() functions in MATLAB. Each histogram had 28 bins and the neighbourhood size was set to 81 pixels.

4.2.3.3 Impact of the number of bins

A 1000×1000 random integer matrix with elements varying from one to the number of bins considered was generated. The number of bins was varied from 1 to 301 and the neighbourhood size was set to 81 pixels. Time to construct one local feature histogram per matrix element was determined using the tic()/toc() functions in MATLAB.

4.2.4 Random forest classifier

A random forest classifier was used to generate all the results presented in this chapter. A third-party MATLAB implementation was used[5]. Unless specified, classification was based on 20 trees. The number of features to sample at each node (*mtry*) was set to the square root of the total number of features.

4.2.5 Segmentation process

4.2.5.1 Classifier training

Training set consisted of raw PCM images with their corresponding ground truths (images manually annotated by human experts). The ground truth images were either binary (foreground in white, background in black) or RGB (black for non-annotated regions, green for foreground, red for background). For each image, N_T pixels were randomly sampled per image for training purposes. Features were computed and stored in a $N_T \times M$ feature matrix where M was the number of features. The classifier was then trained using the feature matrix.

4.2.5.2 Classification

For each image, a feature matrix of dimensions $N \times M$ was computed where N was the number of pixels of the image and M the number of features. The features were given to a previously trained classifier. The result was a vector of labels that was then re-structured to have the same dimensions than the input image. The final output was a binary image whose elements were equal to 1 for foreground objects, 0 otherwise.

4.2.6 Evaluation of segmentation performance

Segmentation performance was evaluated using leave-one-out cross validation (LOOCV). Given a set containing N images, a classifier was trained on N-1 images and tested on the remaining one. This processed was repeated N times until each image was tested. Unless otherwise specified, 50,000 randomly sampled pixels across all images (except the left-out image) were used for classifier training. The segmentation performance was assessed using the F-score as described in section 2.2.4. Two separate image sets were used for mouse and Human embryonic stem cells

[5] https://code.google.com/p/randomforest-matlab/ (last accessed 03.06.2014)

4.2.6.1 Mouse embryonic stem cells images dataset

The image dataset used was the one previously described in section 2.2.3. It was made up of 50 PCM images of mouse embryonic stem cells (mESCs) representative of various stages of a culture (e.g. from seeding to full confluence). The dimension of those images was 250×250 pixels. The ground truths were in the form of a binary mask where foreground (i.e. cell) pixels were annotated in white (RGB value of 255, 255, 255) and background pixels in black (RGB value of 0,0,0).

4.2.6.2 Human embryonic stem cells images dataset

Twenty Human embryonic stem cells (hESC) PCM images were used for evaluation of segmentation performance (Figure 4-14). The dimensions of the images were 500×500 pixels. These images were acquired by Rhys Macown (Department of Biochemical Engineering, UCL). The ground truths were in form of a binary mask where foreground (i.e. hESCs) pixels were annotated in white (RGB value of 255, 255, 255) and background pixels in black (RGB value of 0,0,0).

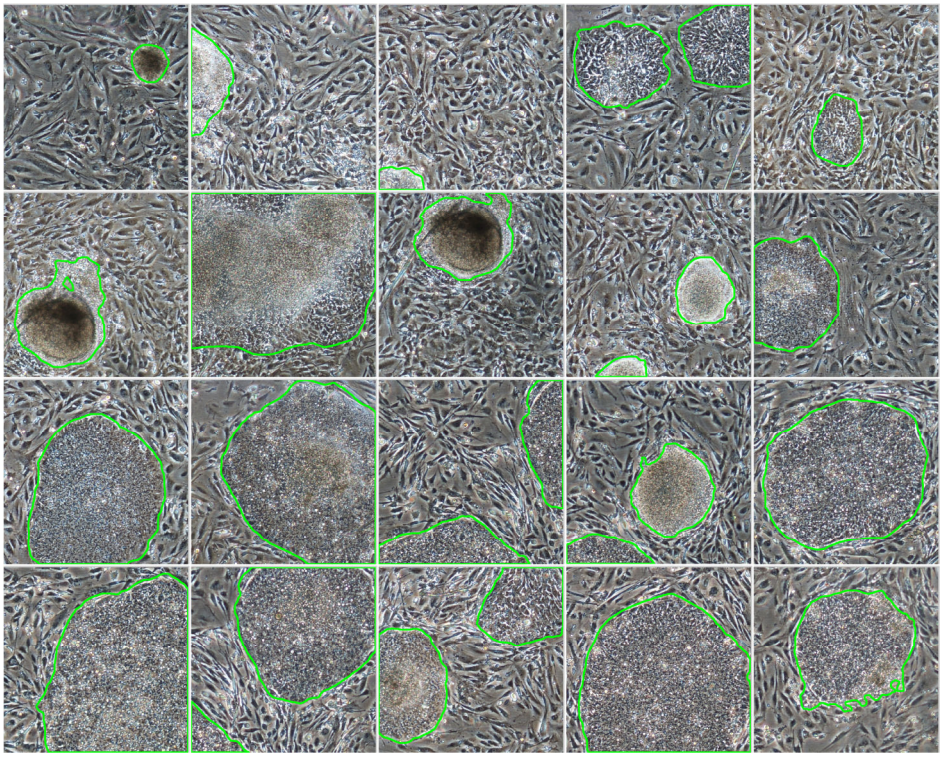


Figure 4-14 Image set used for the evaluation of Human embryonic stem cell (hESC) segmentation performance. The green outlines are the manual annotation of the hESC colonies.

4.3 Results and discussion

In the proposed trainable segmentation scheme, pixels are classified as foreground and background by a random forest classifier based on multi-scale local feature histograms (Figure

4-15). This approach will be thoroughly characterised and its performance evaluated on various PCM images datasets.

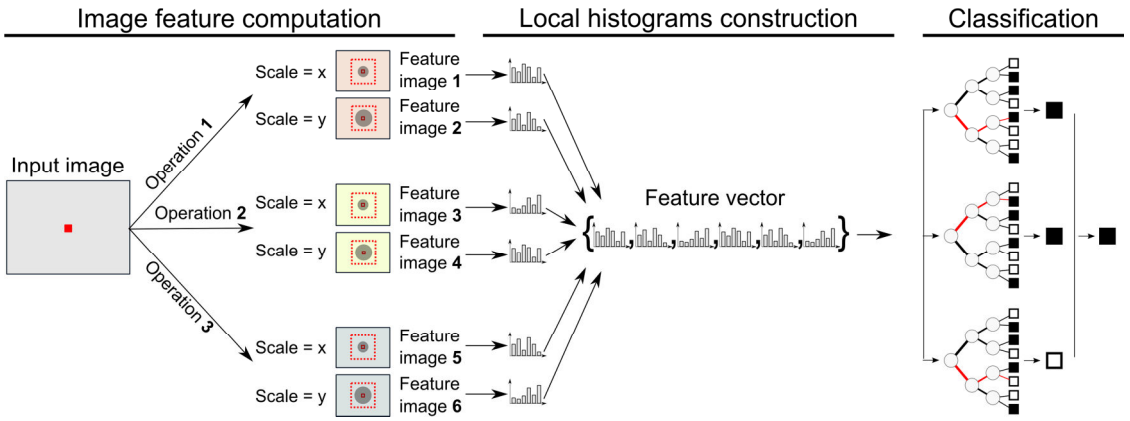


Figure 4-15 Proposed trainable segmentation scheme. Image features are first computed, possibly at different scales. For each combination of features and scales, histograms of feature values within a window centred on each pixel are computed. The local histograms across all combinations are concatenated to form pixel feature vectors, which are then used to predict the pixel class label (foreground or background, here shown as white and black squares, respectively) using a random forest classifier.

Firstly, methods for the construction of local feature histograms are assessed (section 4.3.1). Segmentation performance for cell versus background problems are then evaluated using the mESC PCM image dataset that was employed for the assessment of contrast thresholding algorithms (section 4.3.2). Performance for the more complex task of discriminating between different cell types is evaluated using a PCM image dataset of hESCs co-cultured with mouse embryonic fibroblasts (section 4.3.3). Finally, a high level assessment of the method is given, including one of its ability to alleviate issues related to structured background noise and highly variable cell visual features (section 4.4).

4.3.1 Efficient computation of local image feature histograms

Local feature histogram construction should be as efficient as possible so that it does not become a performance bottleneck. In general, computational complexity for local histogram construction depends on the size of the image, the number of bins, and the size of the neighbourhood considered (Wei and Tao, 2010). As processing speed was of great importance for this project, it was therefore of interest to compare different methods for local feature histogram construction and assess their performance.

Two types of local windows were considered for histogram construction: hard-edged (square) and soft-edged (fuzzy). For the former, frequencies of features were computed using uniform weighting for all pixels regardless of their location within the window considered. In contrast, soft-edged windows used a Gaussian kernel centred at the pixel of interest where the weight of the pixels (i.e. their contribution to the corresponding histogram bin value) decreased as the distance from the central pixel increased. The idea behind this approach was that intuitively,

neighbouring pixels closer to the central pixel should provide information that is more relevant than those situated near the edges of the window.

Three construction methods were considered: sliding window, integral histogram and convolution. Sliding windows were the most straightforward approach to the construction of local feature histograms. A window of a given size was centred at each pixel of the feature image. The values of all pixels within that window were used for the construction of the local feature histogram. This was implemented efficiently using optimised functions in MATLAB (see 4.2.2.1).

Integral histograms were recently described as a very quick and effective approach for the construction of image histograms (Porikli, 2005). In short, an integral histogram was constructed iteratively so that the histogram at a given location of the image encoded all histograms between that location and the origin of the image. The histogram for the local neighbourhood of the pixel of interest can then be quickly computed by simple vector arithmetic on the integral image. The main advantage of this approach was that its computational complexity was constant for a given feature image size, regardless of the size of the neighbourhood considered.

The last construction method considered was based on convolution. Binary masks were created for each bin of the histogram (where elements within a bin were set to 1, zero otherwise), before being convolved with a kernel of size equal to that of the neighbourhood of interest. The histogram for a given pixel was then obtained by concatenating the values found at its location in all convolution outputs. Hard-edged histograms were computed using a kernel whose elements were set to one, while soft-edged histograms were constructed using a Gaussian kernel instead.

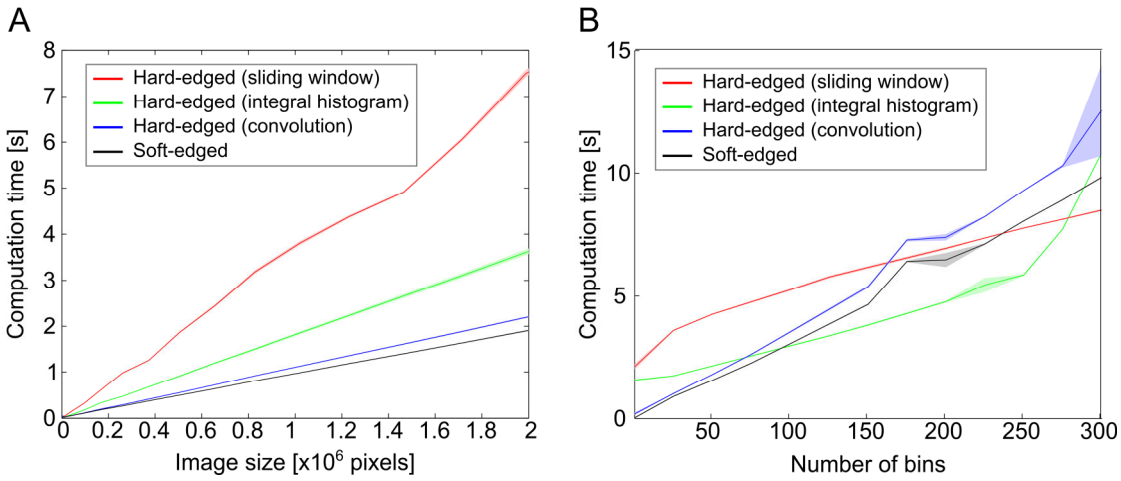


Figure 4-16 Impact of image size (A) and number of bins (B) on histogram computation time. Computation time determination was carried out five times. The lines show the mean and the shaded are of the corresponding colour the standard deviation.

The three construction methods yielded identical results for hard-edged histograms given identical inputs. The criterion of selection for the optimal method was therefore computation speed. The first parameter considered was the size of the input image. For a fixed number of bins (28) and neighbourhood size (81 pixels), the computation time increased nearly linearly with the image size in all cases (Figure 4-16.A). The sliding-window approach was the worst performing, with about 4 seconds required for 1.2 million pixel images (standard size for microscopy images). Integral histograms required less than half this time (~2 seconds) for the same image size. The convolution approach was the best performing with about a second per 1.2 MP image.

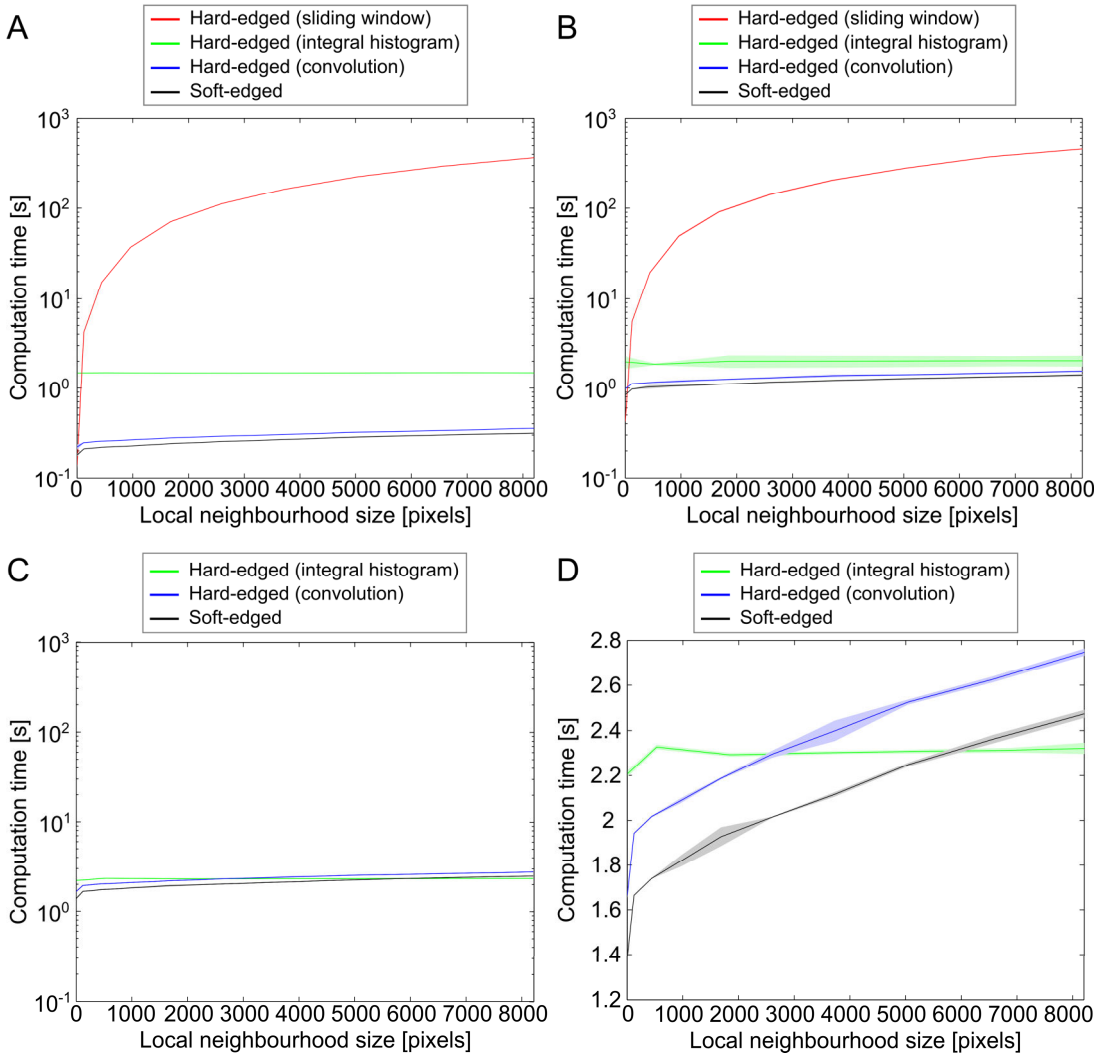


Figure 4-17 Impact of local neighbourhood size on histogram computation time for (A) 7 bins, (B) 28 bins and (C) 56 bins histograms. For A-C the y-axis is a logarithmic scale. (D) Shows the same data than (C) but on a linear scale. Computation time determination was carried out 5 times. The lines show the mean and the shaded are of the corresponding colour the standard deviation.

Next, the impact of the number of histogram bins on the computational time was investigated. This was of importance, as some features required a large number of bins to be properly encoded. The computation time required for the generation of local histograms for a 1000×1000 images (i.e. a million histograms in total) with a neighbour size of 81 pixels was determined for the different construction methods. For low number of bins (<50), the convolution method was the most efficient (Figure 4-16.B). However, as each additional bin required one more convolution, the processing time increased with the number of bins. While it had a higher baseline computation time, for histograms with more than 100 bins the integral histogram approach was the best performing. From 250 bins up, integral histogram computation time soared, most likely due to memory-related bottlenecks. Overall, the sliding-window approach was the slowest and had a higher baseline computation time. It only became the best performing for number of bins above 280. However, such large histograms were unlikely to be considered for this work as they would produce additional bottlenecks at other stages of the segmentation process.

Finally, the computation time for the construction of a million histograms (1000×1000 pixels image) of varying neighbourhood sizes was determined. When considering only seven bins, the convolution method was clearly the best performing with a very low baseline processing time of less than 500 milliseconds, which remained mostly unchanged for higher neighbourhood sizes (Figure 4-17.A). Similarly, integral histograms had a baseline computation time of just above one second, which also remained constant for larger neighbourhoods. The computation time of the sliding-window method increased markedly with the number of pixels in the neighbourhood, before levelling out. As the number of bins was increased to 28, the baseline computation time for the convolution approach increased to about a second, and was very similar to that of the integral histogram method which remained essentially unchanged when compared to the case with 7 bins (Figure 4-17.B). When computing local histogram with 56 bins, the convolution and integral histogram approaches fared very similarly (Figure 4-17.C). The sliding window approach was not included in this last comparison due to excessive computation times. Due to a lower baseline, convolution performed better than integral histograms for smaller neighbourhood sizes (<3000 pixels).

The performance for soft-edged histogram construction by convolution was similar to that obtained for hard-edged histogram construction by convolution. In most cases, the soft-edged computation was slightly more efficient due to the highly optimised implementation of convolution by a Gaussian kernel in MATLAB. Overall, the convolution and integral histogram approaches clearly out-performed the sliding-window method. The former was found to be more suitable for lower number of bins while the latter usually performed better for large histograms (>50 bins).

4.3.2 Trainable segmentation of mouse embryonic stem cells PCM images

The trainable segmentation approach was applied to the mouse embryonic stem cells (mESCs) phase contrast microscopy (PCM) images that have been used to validate the contrast thresholding algorithms in section 2.3. The image features considered for this investigation included raw grayscale pixel intensities, pixel intensities after application of a local contrast filter, and Basic Image Features (Figure 4-18). Each type of feature was first evaluated individually while a subsequent sub-section focuses on their combination.

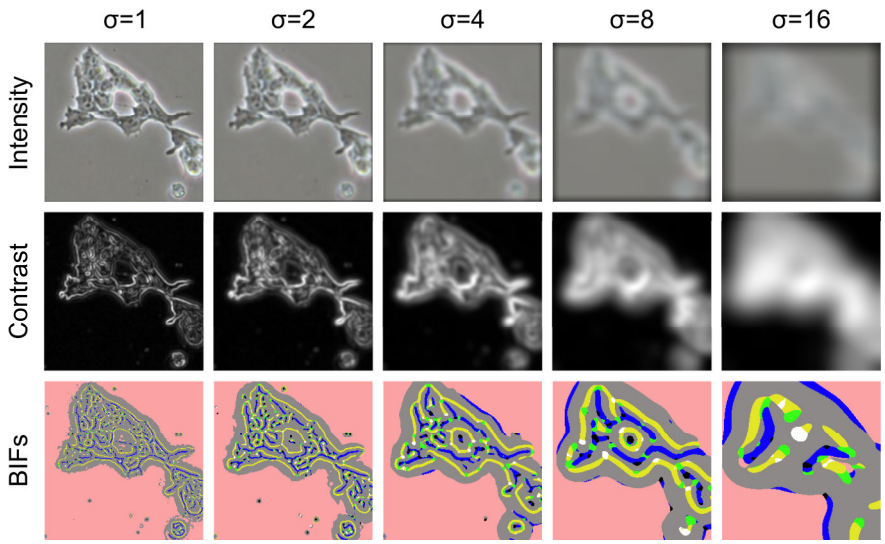


Figure 4-18 Examples of features used for trainable segmentation of mESC PCM images. Each feature (intensity, contrast and Basic Image Features) was computed at 5 different scales (σ). Intensity images were blurred using Gaussian filters of standard deviation equal to the scale shown. Similarly, contrast filters were computed using the scale shown. For Basic Image Features, the scale correspond to the parameter σ, while all images were computed with ε=0.03. Colour codes are shown in Table 4-1.

Segmentation performance was assessed using leave-one-out cross-validation (LOOCV). In short, 50'000 pixels were randomly chosen across 49 manually annotated PCM mESC images to train a random forest classifier. The resulting classifier was then used to predict the labels (i.e. background or cell) for each pixel of a 50th image. Pixel-wise agreement between the predicted labels and manual annotations by human experts was determined using the F-score (see section 2.2.3). This was repeated 50 times so that each image was used for prediction once. The resulting F-scores are indicators of performance for unseen images and thus of the generalisation of the method (to similar images).

For each feature, or combination thereof, three approaches were compared: hard-edged local histograms (heLH), soft-edged local histograms (seLH) and histogram statistics (STA). Both heLH and seLH capture local features by computation of histograms for each pixel of the input image (scaled-down 10-bin histograms for raw intensity and contrast features, full 7-bin histogram for BIFs) whereas STA used statistics (e.g. mean, kurtosis, skewness) to describe full-sized histograms (256 bins for raw intensity and contrast features, 7-bins for BIFs).

The image features for classification were computed at multiple scales (1, 2, 4, 8, and 16). The exact definition of scales was feature-dependent and will be discussed for each feature in their respective sections below. All combinations of those scales were considered (31 in total). Results are shown as leave-one-out cross-validation (LOOCV) F-scores in function of the local window width. As it was not feasible to show the results for all scale combinations and window widths considered, only the best performing scale combination is shown for each local window width. Widths of 0, 5, 10, 15, 20, 30, 45, 100, and 200 pixels were considered. A width of zero signifies that only a single value was used per pixel and per scale instead of histograms (corresponds to the scheme shown in Figure 4-8).

Also of interest was how well the different combinations of scales performed across the range of local window widths tested. The performance rank of each combination for all 9 window widths was recorded. This allowed determining how a given combination of scales generally ranked. Ideally, a combination would rank very highly in most cases, regardless of the window width considered (reduced sensitivity to one of the parameters).

4.3.2.1 Raw pixel intensity features

First, trainable segmentation based on the pixel intensity values of the raw PCM images was assessed. The different scales were obtained by blurring the input PCM image by convolution with a Gaussian kernel of standard deviation corresponding to the scales considered. Ideally, local intensity values would be encoded using 256-bin histograms, or a bin for each possible intensity level of an 8-bit digital image. However, the resulting feature vector for images with N pixels would be of dimensions $N \times 256 \times N_s$, where N_s is the number of scales considered. Even when a single scale was used, these dimensions were unpractical, and resulted in unstainable degradation of the processing time at both training and prediction stages. Instead, scaled-down 10-bin interval histograms were used as a compromise, where the range of possible intensity values was split into evenly spaced bins. The dimensions of the feature vectors were thus equal to $N \times 10 \times N_s$, which was much more manageable.

Table 4-2 Results for trainable segmentation of mESC PCM images based on raw pixel intensity values

Local feature encoding approach	Result type	Window Size	Scales	LOOCV F-score [mean ± std]
Hard-edged 10-bin histogram (heLH)	Worst width	200	4+8+16	0.50 ± 0.33
	Best width	20	1	0.77 ± 0.15
Soft-edged 10-bin histogram (seLH)	Worst width	200	4+16	0.49 ± 0.34
	Best width	30	1+4	0.83 ± 0.12
Statistics of 256-bin histogram (STA)	Worst width	200	1+8+16	0.47 ± 0.33
	Best width	15	1+2+4+8	0.85 ± 0.11

Segmentation performance using 10-bin hard-edged local histograms (heLH, Figure 4-19.A,B), and soft-edged local histograms (seLH, Figure 4-19.C,D) was evaluated. Alternatively, using statistics to describe 256-bin local histograms allowed reducing the size of the feature vector to just $N \times 5 \times N_s$ (STA, Figure 4-19.E and F). A summary of the results obtained for each approach is presented in Table 4-2.

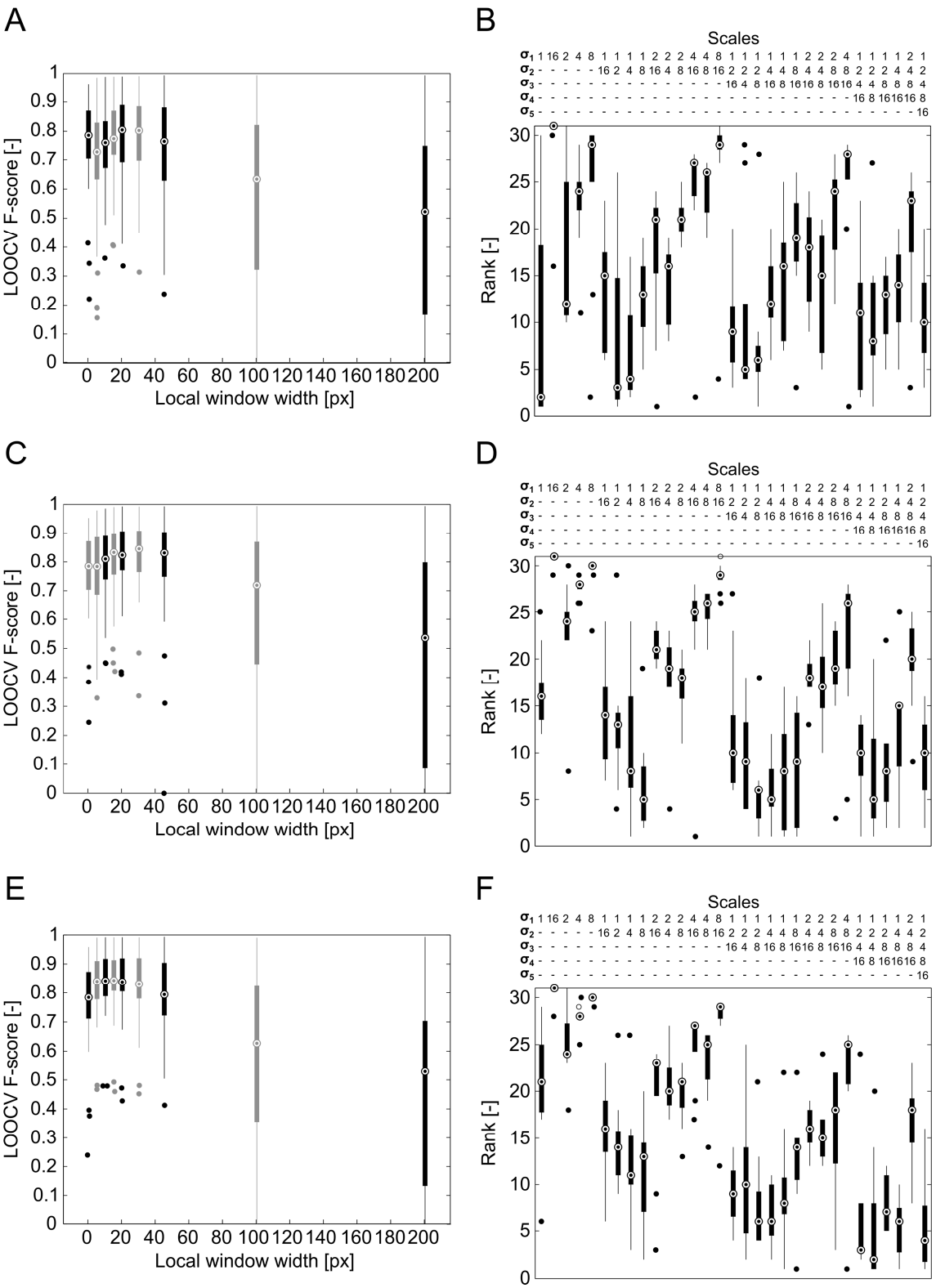


Figure 4-19 Trainable segmentation of mESC PCM images based on intensity features. Hard-edged 10-bin pixel intensity values local histograms (first row, panels A,B), soft-edged 10-bin local pixel intensity values histograms (second row, panels C,D), and statistics of 256-bin local pixel intensity values histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

In all cases, the window width had a clear impact on the segmentation performance. The worst results were consistently obtained for the largest window size considered (200 pixels). When using heLH, small window sizes (5 to 30 pixels) yielded worse results than obtained with a window size of zero (i.e. a single intensity value per pixel per scale). For the other approaches (seLH and STA), performance increased with the window size until an optimal width was reached. Those optimal windows were found to be between 15 and 30 pixels wide depending on the approach used. Further increasing the window width resulted in the deterioration of segmentation performance. While seLH outperformed their hard-edged counterparts, the best performing approach was STA, despite using smaller feature vectors (5 statistical features per scale versus 10 features per scale). This indicated that the statistics approach captured local neighbourhood information suitable for the discrimination between background and cell pixels.

Using a combination of all scales considered was not necessarily the best approach. For heLH, the combinations of two scales (i.e. 1+2 and 1+4) performed consistently well regardless of the window width. For seLH, 1+8, 1+4+16, and 1+2+4+8 were all high-ranking scale combinations. Similarly for STA, 1+2+4+8 was also consistently ranking higher than the others. In general, combination of multiples scales fared well as long as one of the scale was equal to 1. The use of a single scale almost always resulted in poor segmentation performance.

4.3.2.2 Contrast features

As previously shown, the use of contrast filters can greatly improve segmentation performance in the case of conventional image processing approaches (see section 2.3.1). It was thus of interest to determine whether this would also hold true for trainable segmentation. The methodology for this investigation was identical to that employed for raw pixel intensity features evaluation (section 4.3.2.1). The scale of the contrast features was set by varying the scale parameter (σ) of the contrast filter (see section 2.1.2.3).

The trends observed for the contrast features approximately followed those seen above for raw PCM image pixel intensity values (Figure 4-20 and Table 4-3). Large window widths were the worst performers in all cases. When using heLH, a width of 5 pixels resulted in lower performance than when using a single value per pixel (width of 0). However, this was an exception, as in general performance was found to increase with window width until optimal values were reached. Optimal windows were 20 (heLH and STA) and 30 (seLH) pixels wide. Increasing width beyond those optimal values resulted in a rapid and significant decrease in performance. When using contrast features, each of the three approaches yielded results clearly superior to those obtained when using raw pixel intensity values.

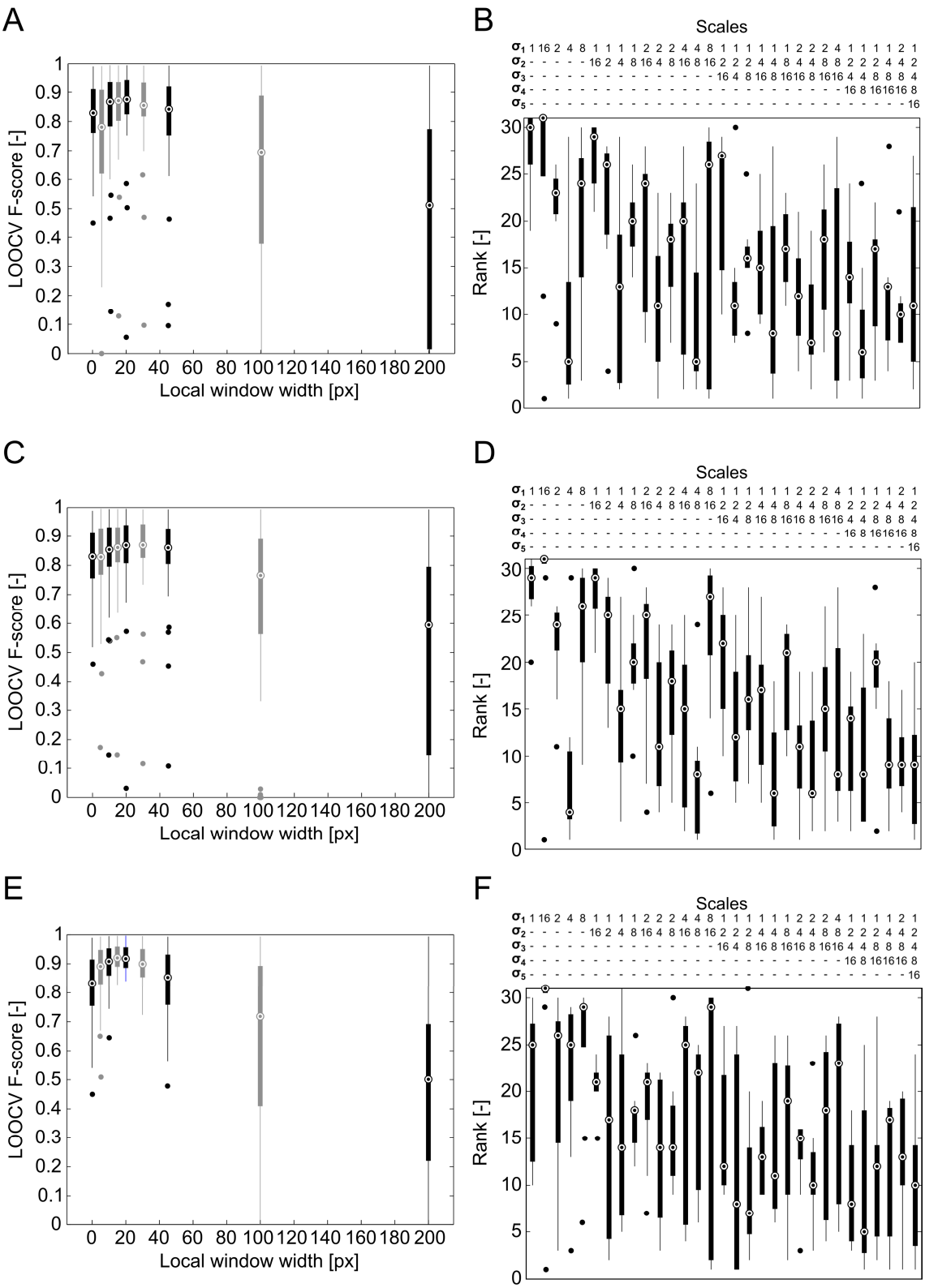


Figure 4-20 Trainable segmentation of mESC PCM images based on contrast filter features. Hard-edged 10-bin local contrast histograms (first row, panels A,B), soft-edged 10-bin local local contrast histograms (second row, panels C,D), and statistics of 256-bin local contrast values histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

STA performed very well and outclassed other methods with LOOCV F-scores of 0.92 ± 0.04 obtained using a combination of three scales (1+2+4) and a window width of 20 pixels. This could indicate that the 10-bin interval histograms were not suitable to capture contrast-filtered image information. Indeed, it might be that only a few bins of said histograms were having a significant effect on the pixel classification outcome. To confirm this, we exploited the ability of the random forest classifier to generate importance measurements for each feature used for classification in the form of the mean decrease in Gini index (i.e. measure of how much a feature contributed to the class separation during classification). For heLH (10 bins), more than half the features were found to have no measurable effect on the classification outcome (Figure 4-21.A). In contrast, all bins had some importance when using STA (Figure 4-21.B). This indicated that despite the lower number of features used per scale, STA provided information to the classifier that was more suitable for the discrimination between background and cell pixels.

Table 4-3 Results for trainable segmentation of mESC PCM images based on contrast features

Local feature encoding approach	Result type	Window width [pixels]	Optimal scales	LOOCV F-score [mean ± std]
Hard-edged 10-bin histogram (heLH)	Worst width	200	16	0.45 ± 0.37
	Best width	20	1+4+8	0.85 ± 0.15
Soft-edged 10-bin histogram (seLH)	Worst width	200	16	0.52 ± 0.35
	Best width	30	1+4+8	0.85 ± 0.15
Statistics of 256-bin histogram (STA)	Worst width	200	16	0.49 ± 0.30
	Best width	20	1+2+4	0.92 ± 0.04

When using either heLH or seLH, scale combinations generally performed well if they at least included a scale equals to 4. This is in stark contrast with the raw pixel intensity features results where at least one of the scales being equal to 1 was required to obtain satisfactory segmentation performance. For STA, combining 3 or more scale resulted in good performance, the 1+2+4+8 combination being consistently high ranking.

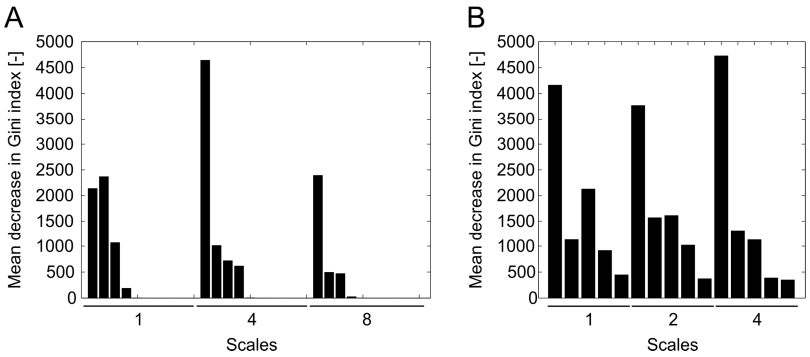


Figure 4-21 Features importance for segmentation using contrast features. (A) Hard-edged 10-bin local histograms (10 features per scale) and (B) statistics of 256-bin local histograms (5 features per scale). The mean decrease in Gini index is a measure of the contribution of a given feature to the separation of the classes during classification.

4.3.2.3 Basic Image Features

In addition to raw pixel values and local contrast features, it was of interest to determine if a set of higher level features, such as texture descriptors, could lead to an increase in segmentation performance. Basic Image Features (BIFs) are one such descriptor. BIFs classify all pixels of an image as one of seven categories according to local symmetries and structure. They were expected to allow classification of pixels based on a higher level of abstraction. They are also particularly adequate for trainable segmentation as local BIFs histograms have the dimensions $N \times 7 \times N_s$ (where N and N_s are the number of pixels of an image and the number of scales considered, respectively). Results for the three histogram constructions approaches (heLH, seLH, and STA) are presented in Figure 4-22 and summarised in Table 4-4.

Table 4-4 Results for trainable segmentation of mESC PCM images using Basic Image Features

Local feature encoding approach	Result type	Window width [pixels]	Optimal scales	LOOCV F-score [mean ± std]
Hard-edged 7-bin histogram (heLH)	Worst width	200	16	0.48 ± 0.32
	Best width	15	1+4	0.92 ± 0.05
Soft-edged 7-bin histogram (seLH)	Worst width	200	8	0.50 ± 0.34
	Best width	20	1+2+8	0.92 ± 0.05
Statistics of 256-bin histogram (STA)	Worst width	200	16	0.49 ± 0.28
	Best width	20	1+2+4+8	0.91 ± 0.05

For the three feature encoding approaches, the segmentation performance increased with the width of the window until it peaked between 15 and 20 pixels. For widths beyond that optimal value, performance was significantly degraded, reaching a minimum at 200 pixels wide (Figure 4-22). heLH and seLH produced identical optimal results with LOOCV F-scores of 0.92 ± 0.05. While STA resulted in better segmentation performance than heLH and seLH for pixel intensity and local contrast features, it produced slightly lower LOOCV F-scores when using BIFs. This could be explained by the fact that BIFs heLH and seLH only had 7 bins, which captured the entirety of the BIFs images information. In contrast, intensity and local contrast heLH and seLH had 10 bins to describe 256 possible values, leading to an obvious loss of information. In those cases, STA captured more information and was therefore more appropriate as feature vector. When using BIFs, however, both heLH and seLH were found to be more suitable.

Similarly to what was observed for intensity features, scales combinations tended to perform well as long as they included a scale equal to 1. Combinations of at least three scales were also found to be more consistently highly ranking than those of two or less.

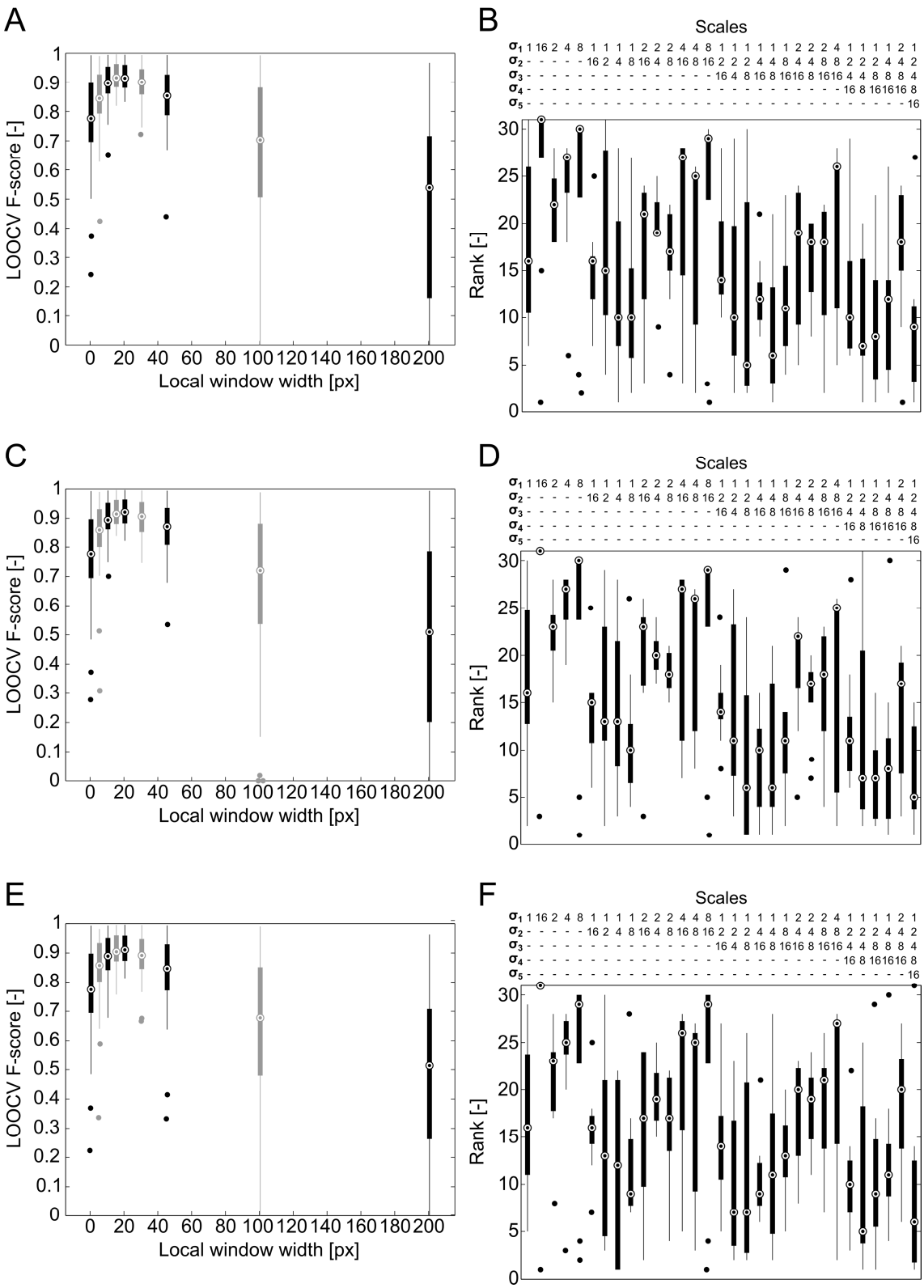


Figure 4-22 Trainable segmentation of mESC PCM images based on Basic Image Features (BIFs). Hard-edged 7-bin basic image features (BIFs) histograms (first row, panels A,B), soft-edged 7-bin local BIFs histograms (second row, panels C,D), and statistics of 7-bin local BIFs histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

4.3.2.4 Combination of features

Segmentation performance was evaluated for the case where multiple types of features were used together. In order to minimise the computational time required generating these results, a single combination of scale was investigated (1+2+4+8+16). Raw intensity, contrast and Basic Image Features computed in local windows of varying widths were used to train a random forest classifier (Figure 4-23).

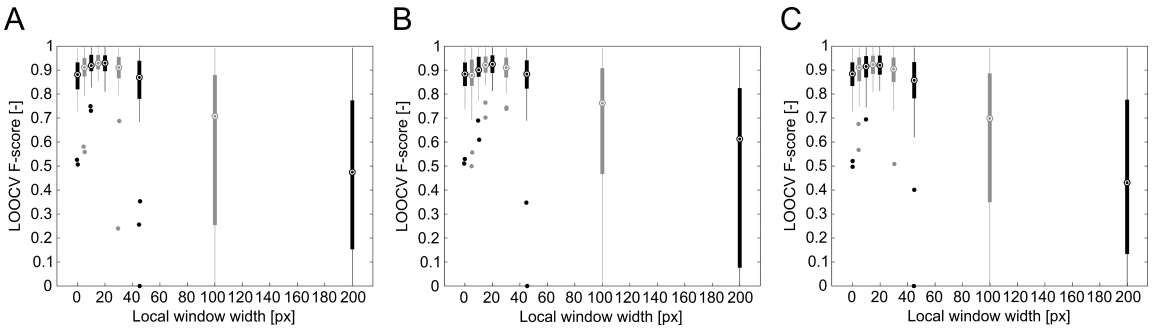


Figure 4-23 Trainable segmentation using combinations of features. (A) 10-bin hard-edged local histograms (heLH), (B) 10-bin soft-edged local histograms (seLH), and (C) statistics of 256-bin hard-edged local histograms (STA). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

The best segmentation performance obtained for each approach was LOOCV F-scores of 0.93 ± 0.04 (window width of 15 pixels), 0.92 ± 0.05 (window width of 20 pixels), and 0.92 ± 0.05 (window width of 20 pixels) for heLH, seLH, and STA respectively. This was only a slight improvement over the best score obtained using BIFs alone. The importance of each feature for classification for the heLH case was determined using the mean Gini decrease, a measure of the contribution of a given feature to the separation of classes during classification (Figure 4-24).

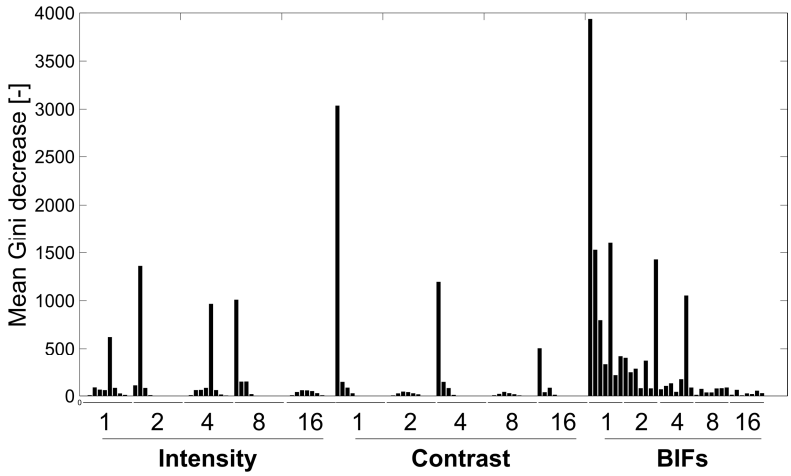


Figure 4-24 Feature importance when combining raw intensity, contrast and Basic Image Features at five different scales. The mean decrease in Gini index is a measure of the contribution of a given feature to the separation of the classes during classification.

For raw intensity features, all scales were found to have some degree of importance beside the scale equal to 16. Similarly for contrast features, scales equal to 2 and 8 had very low Gini decrease values associated with them. In both cases, many bins of the histogram appeared not to

be important for classification. In contrast, Basic Image Features were of high importance for lower scales (1, 2 and 4), while higher scales had a much smaller impact. However, most bins of the histograms had at least some degree importance, unlike for the other two feature types. Once again, this showed that BIFs were particularly suitable for trainable segmentation approaches based on local feature histograms, as only 7 bins were required to describe image information in its entirety whereas both raw intensity and contrast features used 10-bin histograms to describe 256 possible values, thus inevitably leading to a loss of information.

4.3.2.5 Summary

Table 4-5 Best performing settings for each feature types considered for trainable segmentation of mESC PCM images

Feature type	Approach	Window width [pixels]	Scales	LOOCV F-score [mean ± std]	Processing time[1] [s]
Intensity	STA	15	1+2+4+8	0.85 ± 0.11	3.50
Contrast	STA	20	1+2+4	0.92 ± 0.04	2.73
BIFs	heLH	15	1+4	0.92 ± 0.05	0.07
Combination	heLH	15	1+2+4+8+16	0.93 ± 0.04	0.51

1. As computed for a 250x250 PCM image. Value shown is the mean of 50 computations carried out using an Intel i7-4770K CPU with 16GB of RAM

In the vast majority of cases, using local feature histograms resulted in significantly higher segmentation performance than that obtained based on a single value per pixel per scale (Table 4-5). Moreover, the best results were obtained by combining the various features considered. Contrast features (using statistics of 256-bin local histograms) and Basic Image Features (using a 7-bin hard-edged histogram) resulted in very similar performance. Intensity features produced the worse results amongst the three types of features investigated. When employing statistics of local feature histograms (STA), the processing time soared. Thus, despite slightly lower performance, Basic Image Features appeared to be the better choice due to a significantly lower processing time.

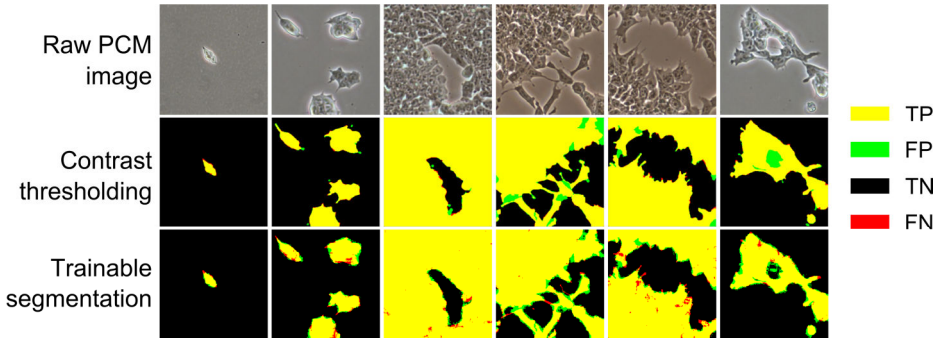


Figure 4-25 Comparison of mESCs PCM images segmentation outputs between contrast thresholding and trainable segmentation. For contrast thresholding, the nseSTDEV method was used with optimal parameters as determined by an extensive grid search (section 2.3.2). For trainable segmentation, hard-edged local histograms of BIFs at scale 1+4 were used. The colour codes are as following: yellow (true positives, TP), green (false positives, FP), black (true negatives, TN) and red (false negatives, FN).

Interestingly, the best trainable segmentation results remained below those of segmentation based on contrast thresholding and post-hoc halo correction, where LOOCV F-scores of 0.95 ± 0.04 were obtained for the same set of images (see section 2.3.2). The difference can be explained by multiple factors. Whereas the contrast filter approach was specifically devised for the segmentation problem at hand, trainable segmentation was by nature a generic approach that did not rely on any *a-priori* information. In this regard, it performed remarkably well. The main difference observable when comparing segmentation outputs is the tendency of trainable segmentation to produce noisier outputs, most likely due to the lack of post-processing such as hole filling or small objects removal (Figure 4-25). In most cases, the halo artefacts were successfully corrected by the trainable segmentation approach. Another interesting aspect is the suitability of the manual annotations used as ground truths. The random forest classifier might be picking up 'bad habits' from erroneous ground truth data. In contrast, a more conventional approach such as contrast thresholding would not suffer as much from this issue. It would therefore be interesting to devise methods for the generation of ground truth data that would not rely on the subjectivity of a human expert.

In summary, trainable segmentation was used to produce high quality segmentation of mESC PCM images. The short time required for the training of the classifier more than made up for segmentation performance that was slightly lower when compared to the contrast thresholding approach. In addition, trainable segmentation is expected to be much more versatile, as it can potentially deal with more complex segmentation scenarios. This is further investigated for Human embryonic stem cell PCM images segmentation below (4.3.3) as well as in chapter 5 where trainable segmentation is applied to the monitoring of cultures in a microfabricated culture device.

4.3.3 Trainable segmentation of Human embryonic stem cell PCM images

Phase contrast microscopy (PCM) images of human embryonic stem cell (hESC) are complex in nature. Indeed, hESC are usually co-cultured together with feeder cells, whose role is to provide factors that are essential for the growth and pluripotency maintenance of the hESC colonies. PCM images of hESC will thus also contain feeder cells, such as mouse embryonic fibroblasts (MEFs) for this particular experimental setup (Figure 4-26.A). Unlike the segmentation of mouse embryonic stem cells (mESC) PCM images that involves the discrimination between foreground objects (cells) and background, hESC image segmentation involves the discrimination between two types of foreground objects (hESC and MEFs), and background (Figure 4-26.B). Though, it remained a two classes classification problem (Figure 4-26.C), with the positive class being hESC colonies and the negative class corresponding to the rest of the image (background and MEFs).

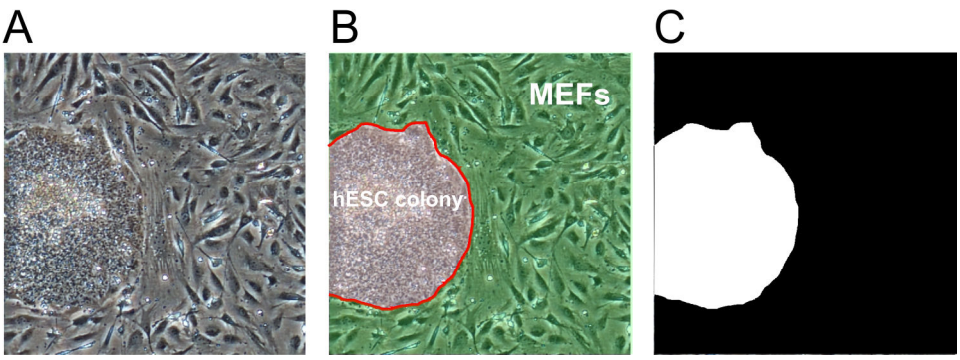


Figure 4-26 Segmentation of hESC PCM images. (A) Raw PCM images of a hESC colony on feeder background. (B) Image from A with overlays indicating the separation between the hESC colony (red border and transparent overlay) and the rest of the image made up of mouse embryonic fibroblasts and background (green overlay). (C) Expected segmentation outcome with the hESC colony as the positive class (white) and the rest of the image as negative class (black).

As was the case for trainable segmentation of mESC PCM images, three approaches for the encoding of local feature information were investigated: hard-edged local histograms (heLH), soft-edged local histograms (seLH), and statistics of local histograms (STA). The type of features evaluated was also the same: raw pixel intensity, pixel intensity after application of a contrast filter, and Basic Image Features (Figure 4-27).

Again, performance evaluation was carried out based on the F-score, a measurement of the agreement between the output of the algorithm and an image manually annotated by a human expert (see section 2.2.4). The training set was made up of 20 500×500 pixels hESC PCM images. For all results shown, leave-one-out cross-validation (LOOCV) was employed. To evaluate the generalisation of the method to unseen images, all images but one (i.e. 19 images) were used to train a random forest classifier, which was then in turn used to classify the pixels of the image that was left out. This processed was repeated 20 times, so that each image of the training set was left out once. A range of local window widths were investigated, from 0 pixel (a single value per pixel per scale) to 400 pixels.

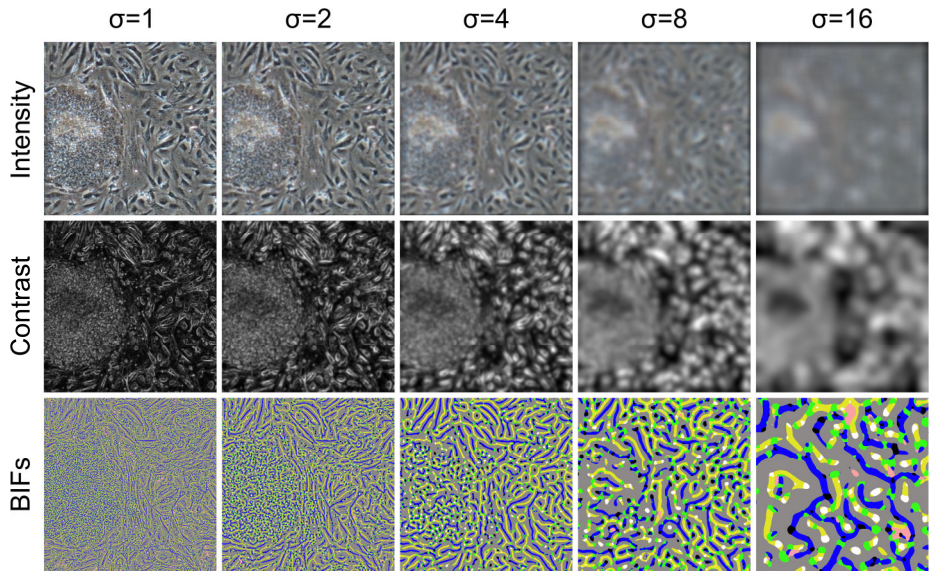


Figure 4-27 Examples of features used for trainable segmentation of hESC PCM images. Each feature (intensity, contrast and BIFs) was computed at 5 different scales.

4.3.3.1 Raw pixel intensity features

The first type of features to be evaluated for hESC trainable segmentation was raw pixel intensity features. The different scales were computed by blurring the image using a Gaussian kernel of standard deviation corresponding to the scale considered. Both hard-edged (heLH) and soft-edged (seLH) local histograms approaches were based on scaled-down 10-bin histograms. Using the 256-bin histograms necessary to accurately represent local pixel intensity information would be unpractical in terms of processing time and memory requirements. Results are presented in Figure 4-28 and summarised in Table 4-6.

Table 4-6 Results for trainable segmentation of hESC PCM images based on raw intensity features

Local feature encoding approach	Result type	Window width	Scales	LOOCV F-Score [mean ± std]
Hard-edged 10-bin histogram (heLH)	Worst width	400	1+4	0.45 ± 0.37
	Best width	100	1+2	0.67 ± 0.24
Soft-edged 10-bin histogram (seLH)	Worst width	400	4+8+16	0.46 ± 0.34
	Best width	100	1+2	0.71 ± 0.24
Statistics of 256-bin histogram (STA)	Worst width	400	1+2+4	0.43 ± 0.35
	Best width	100	1+2+4	0.80 ± 0.11

Overall, the segmentation performance achieved using raw intensity features was poor. Indeed, the best LOOCV F-score obtained was 0.80 ± 0.11 for a window width of 100 pixels using the STA approach. In all cases, the optimum window width was around 100 pixels, and low segmentation performance was observed for the two extremes of the window widths spectrum, 0 pixel (i.e. a single value per pixel per scale) and 400 pixels. The fact that the STA approach resulted in the best performance indicated that the 10-bin histograms used for both heLH and seLH were not adequate for the description of local raw intensity features. In general, scale combinations containing multiple small scales (1, 2 or 4) performed better than those including only larger scales.

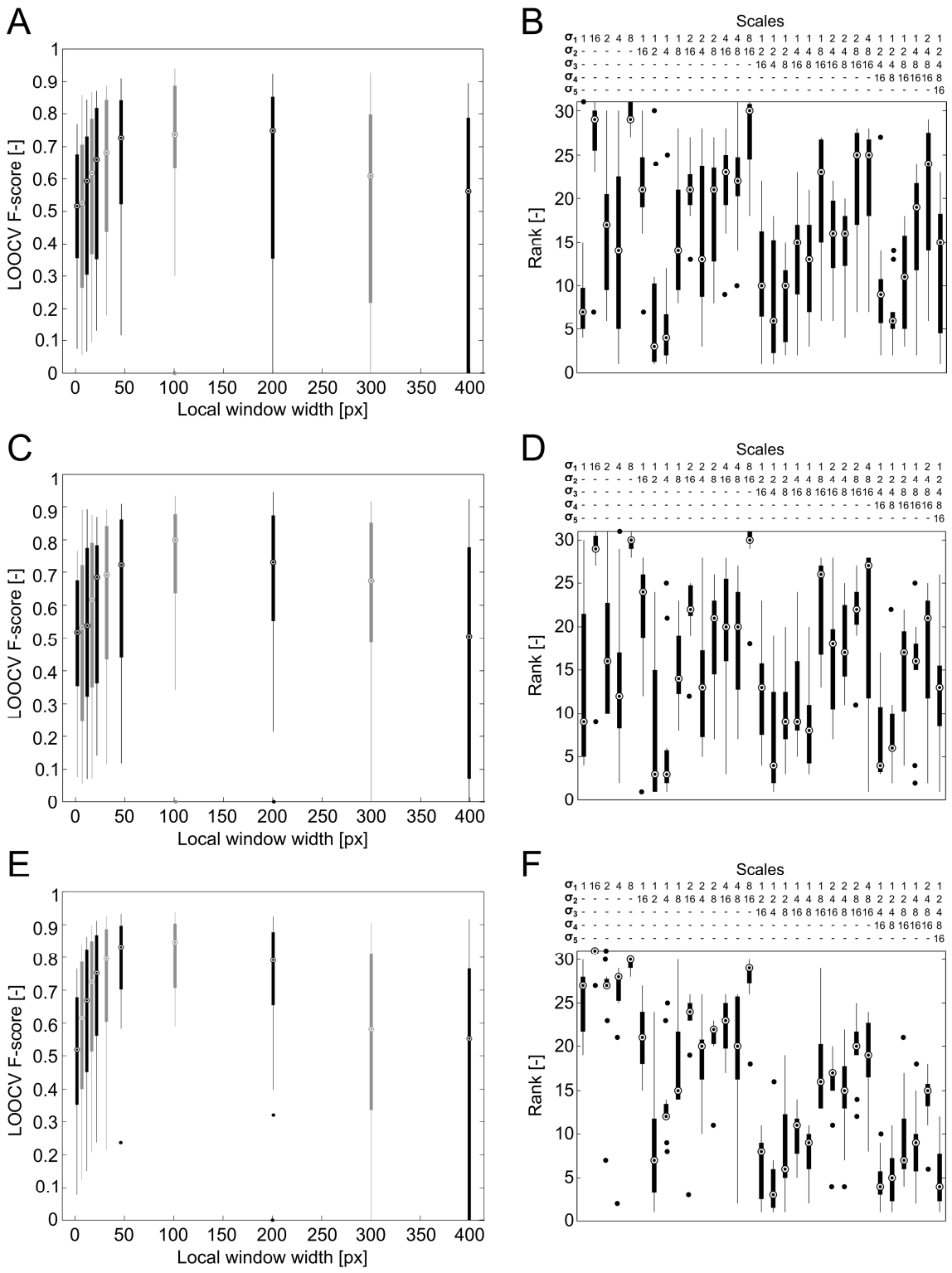


Figure 4-28 Trainable segmentation of hESC PCM images based on pixel intensity values. Hard-edged 10-bin pixel intensity values local histograms (first row, panels A,B), soft-edged 10-bin local pixel intensity values histograms (second row, panels C,D), and statistics of 256-bin local pixel intensity values histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

4.3.3.2 Contrast features

It was investigated whether the output of a contrast filter described earlier (normalised soft-edged standard deviation filter, see section 2.3.1) could be used as feature for the trainable segmentation of hESC PCM images. As for intensity features, both heLH and seLH approaches were limited to down-scaled 10-bins histograms due to processing time and memory constraints. Different scales were computed by varying the scale parameter (σ) of the contrast filter (see section 2.1.2.3). Results are presented in Figure 4-29 and summarised in Table 4-7.

Table 4-7 Results for trainable segmentation of hESC PCM images based on contrast features

Local feature encoding approach	Result type	Window Size	Scales	LOOCV F-Score [mean ± std]
Hard-edged 10-bin histogram (heLH)	Worst width	400	1+8	0.50 ± 0.37
	Best width	100	1+2	0.87 ± 0.10
Soft-edged 10-bin histogram (seLH)	Worst width	5	1+2+4	0.51 ± 0.31
	Best width	100	1+2+8+16	0.87 ± 0.10
Statistics of 256-bin histogram (STA)	Worst width	400	4+16	0.46 ± 0.36
	Best width	100	1+2	0.86 ± 0.10

Similarly to the results obtained for intensity features, performance of the trainable segmentation of hESC PCM images based on contrast features was at its worst for the extreme ends of the range of local window sizes investigated. In general, performance increased together with the window width until the optimal window size was reached. The latter was found to be 100 pixels for all the local feature encoding approaches. Increasing window width beyond the optimal value resulted in the rapid degradation of segmentation performance.

Identical results were obtained using heLH (1+2 scale combination) and seLH (1+2+8+16 combination) with LOOCV F-scores of 0.87 ± 0.10. The STA approach resulted in slightly lower segmentation performance with LOOCV F-scores of 0.86 ± 0.10. These results represented a significant improvement over was what achieved using only raw pixel intensity values (Table 4-6). Again, combinations of scales that included at least some of the smaller scales generally fared better than those only including larger scales. Having one of the scale being equal to 1 was generally a requirement to produce satisfactory segmentation performance.

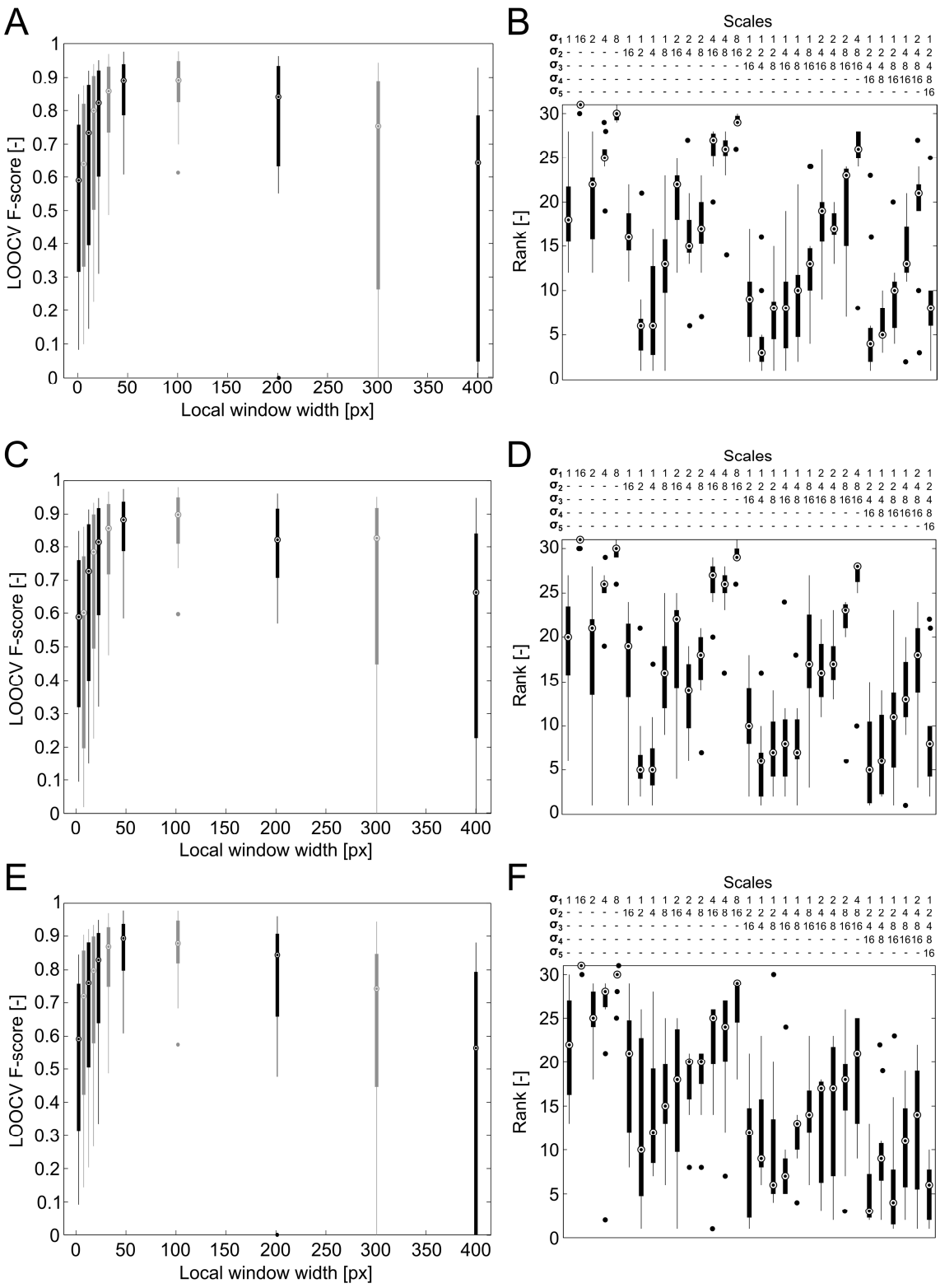


Figure 4-29 Trainable segmentation of hESC PCM images based on contrast features. Hard-edged 10-bin local contrast feature histograms (first row, panels A,B), soft-edged 10-bin local contrast feature histograms (second row, panels C,D), and statistics of 256-bin local contrast feature histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

4.3.3.3 Basic Image Features

It was investigated whether Basic Image Features (BIFs) could be used for the trainable segmentation of hESC PCM images. BIFs are high level features sensitive to structure and texture characteristics of images that are not necessarily captured by either raw intensity or contrast features. BIFs were thus thought to be more suitable for the discrimination between Human embryonic stem cells and mouse embryonic fibroblast pixels, as both cell types exhibit slightly different textures. The scale of the feature was varied by altering the values of the scale parameter (σ) used for the BIFs computation (see section 4.2.1.3). Results are presented in Figure 4-30 and summarised in Table 4-8.

Table 4-8 Trainable segmentation of hESC PCM images based on basic image features

Local feature encoding approach	Result type	Window Size	Scales	LOOCV F-score [mean ± std]
Hard-edged 7-bin histogram (heLH)	Worst width	0	16	0.44 ± 0.22
	Best width	100	1+2+4+8	0.90 ± 0.07
Soft-edged 7-bin histogram (seLH)	Worst width	0	16	0.44 ± 0.22
	Best width	100	1+2+8	0.90 ± 0.07
Statistics of 7-bin histogram (STA)	Worst width	0	16	0.44 ± 0.22
	Best width	100	1+2+8	0.89 ± 0.07

For all local feature encoding approaches tested (heLH, seLH, and STA), the worst and optimal window widths were found to be 0 and 100 pixels, respectively. Performance increased with width until the optimal value was reached. Further increasing width past this optimal value resulted in the deterioration of the segmentation performance. Both heLH (combinations of scales 1+2+4+8) and seLH (combination of scales 1+2+8) produced identical results with LOOCV F-scores of 0.90 ± 0.07. The STA approach resulted in very similar performance, with LOOCV F-scores of 0.89 ± 0.07.

For both heLH and seLH, the performance increased as more scales were combined as long as said combinations included a scale equal to 1. For the STA approach, combining multiple scales was detrimental to the segmentation performance.

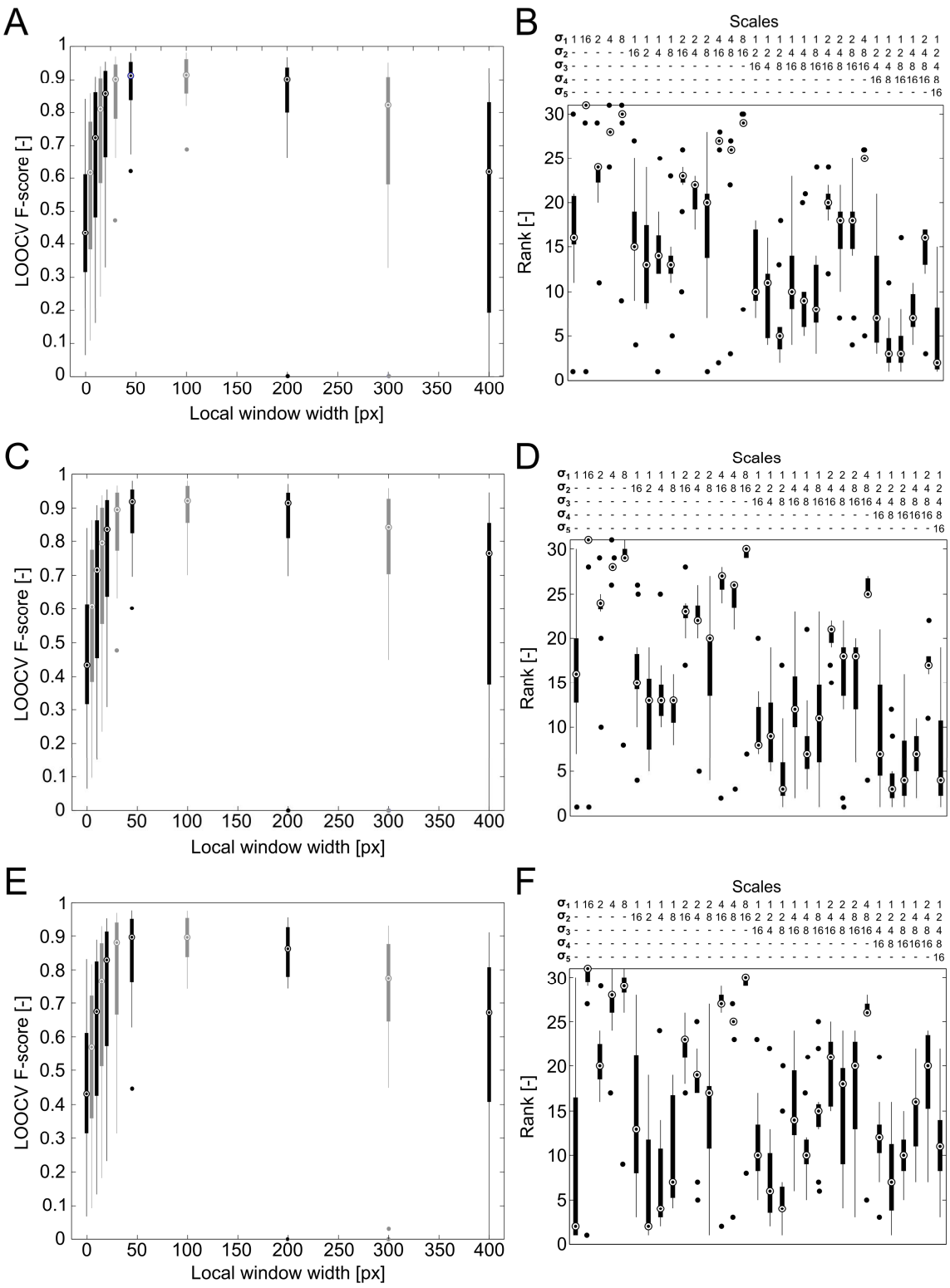


Figure 4-30 Trainable segmentation of hESC PCM images based on Basic Image Features (BIFs). Hard-edged 7-bin basic image features (BIFs) histograms (first row, panels A,B), soft-edged 7-bin local BIFs histograms (second row, panels C,D), and statistics of 7-bin local BIFs histograms (third row, panels E,F). (A,C,E) F-score results after cross-validation in function of the local window width. For each window width, the data shown represent the best performing scale combination. The colour of the data points is alternated for ease of interpretation. (B,D,F) Ranks achieved for each combination of scales (each combination was tested nine times, one for each local window width from A, C and E). In all cases, the circles are the medians, the edges of the boxes are the 25th and 75th percentiles and the whiskers extend to the most extreme data points (not including outliers). The 'o' markers represent outliers, which are values outside of the range [75th quartile - 1.5*(75th quartile-25th quartile); 75th quartile + 1.5*(75th quartile-25th quartile)].

4.3.3.4 Summary

The segmentation performance for the combination of the different features could not be evaluated for hESC PCM images due to the computational complexity resulting from large feature vectors.

When using individual features, Basic Image Features (BIFs) clearly outclassed both intensity and contrast features. In particular, intensity features were associated with very low segmentation performance, indicating that raw pixel intensity might not be sufficient to accurately discriminate between hESC colonies and the background of mouse embryonic fibroblasts. BIFs also fared well in terms of processing time, which was equivalent to contrast features while maintaining significantly higher segmentation performance.

Table 4-9 Best performing settings for each feature types considered for hESC trainable segmentation

Feature type	Approach	Window width [pixels]	Scales	LOOCV F-score [mean ± std]	Processing time[1] [s]
Intensity	STA	100	1+2+4	0.80 ± 0.11	5.27
Contrast	heLH / seLH	100	1+2 / 1+2+8+16	0.87 ± 0.10	0.16 / 0.26
BIFs	heLH / seLH	100	1+2+4+8 / 1+2+8	0.90 ± 0.07	0.22 / 0.13

1. As computed for a 500x500 PCM image. Value shown is the mean of 50 computations carried out using an Intel i7-4770K CPU with 16GB of RAM

4.4 Comparison with other trainable segmentation schemes

The proposed trainable segmentation method based on multi-scale BIFs histograms was compared to Ilastik and the Weka trainable segmentation plugin for FIJI (Hall et al., 2009; Schneider et al., 2012; Sommer et al., 2011b), two software packages that demonstrated early on how trainable segmentation could represent a viable alternative to conventional approaches.

Table 4-10 Comparison of the proposed trainable segmentation scheme (multi-scale BIFs histograms) with schemes implemented in two other trainable segmentation software packages. All results were obtained using equivalent user annotations. For Ilastik, all feature types at all scales were considered. For Weka trainable segmentation, Gaussian blur, Sobel filter, Hessian, difference of Gaussian, and membrane projection features were used with sigma varying from 1.0 to 16.0. Results shown as mean F-score ± std

Scheme	mESC images	hESC images
Ilastik (all features and scales)	0.79 ± 0.16	0.65 ± 0.17
Weka trainable segmentation	0.76 ± 0.18	0.60 ± 0.16
Proposed scheme	0.80 ± 0.14	0.92 ± 0.02

Like the proposed scheme, both Ilastik and Weka use random forest to learn pixel-wise segmentation based on generic image features. The main difference being that Ilastik and Weka use a single value per feature per scale to construct pixel feature vectors whereas the proposed method employs local feature histogram to capture spatial context.

Carrying out a fully objective comparison is challenging. All three implementations rely on interactive annotation of images by the user. In addition, a large number of combinations of features and scales can be used, making a thorough exploration of the space of possible parameters for implementation a very lengthy process. Instead, a semi-quantitative approach was taken. Two sets of images were considered: four mESC images (Figure 4-31) and three hESC images (Figure 4-32). In contrast with the training datasets used previously, which were down-sized, the images used here were full size (1280×960). The user annotations were equivalent for all three implementations, though not identical due to differences in the different software packages. For Ilastik and the proposed scheme, annotations were based on brush strokes of similar size. For the Weka trainable segmentation, however, annotations are limited to thin lines. In this case, brush strokes were mimicked by having multiple parallel lines spanning the width of the strokes (e.g. one at each edge of the stroke and one in the centre). The classifiers were trained using four and three images for the mESC and hESC experiments, respectively.

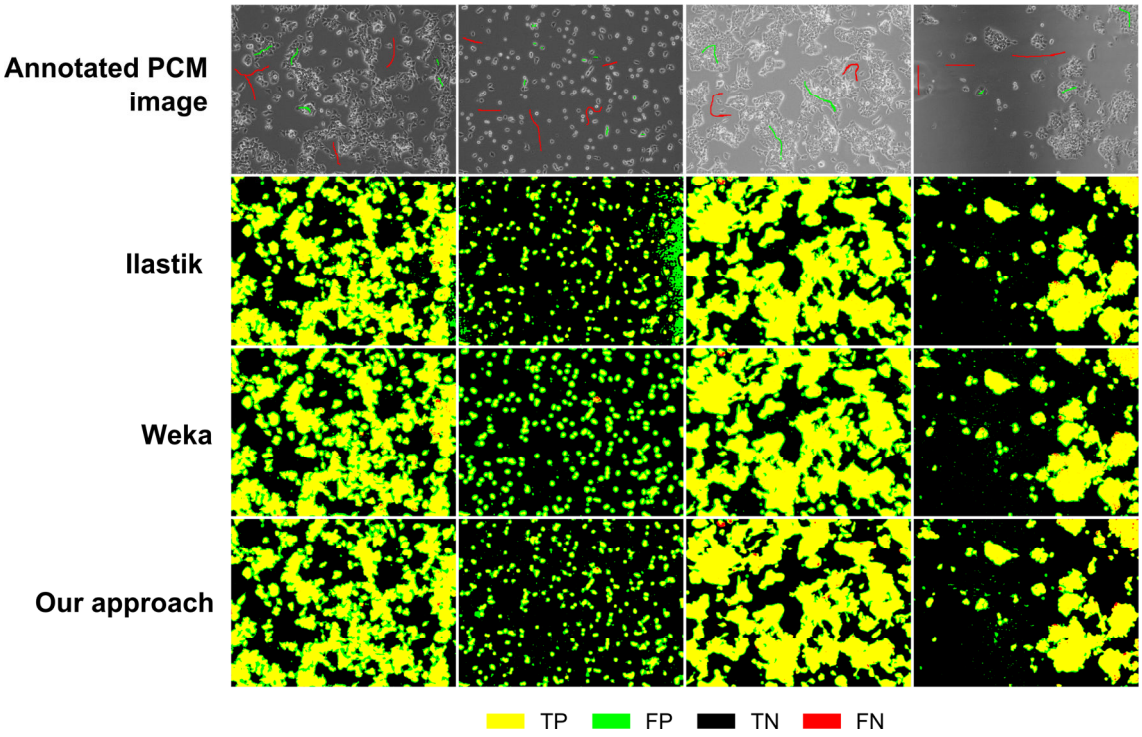


Figure 4-31 Comparison of different trainable segmentation schemes for four full-resolution mESC images. The first row shows the raw PCM image with user annotations, with red and green strokes representing background and foreground, respectively.

For the mESC images, all three schemes yielded similar results (Figure 4-31 and Table 4-10). For the annotations and processing parameters considered, the proposed scheme (multi-scale local BIFs histograms) performed slightly better than the two other implementations. Very few false negatives were observed across the segmentation outputs. Comparatively, the number of false positives was high, usually due to the partial misclassification of halo artefacts. Small cellular objects (as shown in the second column of Figure 4-31) appeared to pose a challenge to all three methods considered whereas larger colonies were detected much more reliably. For the

images considered, the use of local feature histogram only provided a marginal advantage over the other schemes.

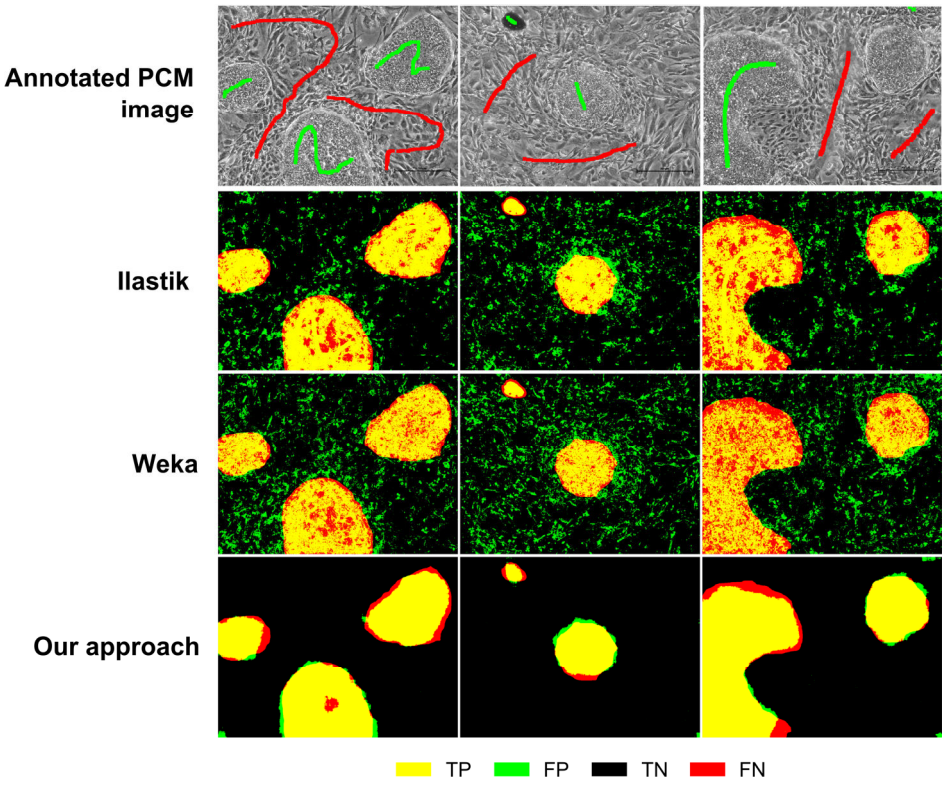


Figure 4-32 Comparison of different trainable segmentation schemes for three full-resolution hESC images. The first row shows the raw PCM image with user annotations, with red and green strokes representing background and foreground, respectively.

Differences in segmentation performance between the three schemes compared were much more marked for hESC images (Figure 4-9 and Table 4-10). Both Ilastik and Weka had issues differentiating between the hESC colonies and the fibroblast background, as shown by the very high amount of noise seen in the segmentation outputs. Despite trying a wide range of feature types and parameters, this issue could not be alleviated to a suitable degree. In contrast, the proposed scheme resulted in very clean segmentation of hESC colonies, as reflected by significantly improved F-scores. In this case, the use of local feature histograms proved to be critical in achieving reasonable segmentation performance and thus represented a significant improvement over previously described methods.

4.5 Conclusion and outlook

Trainable segmentation was shown to be a viable approach for the segmentation of mESC and hESC PCM images. In the case of mESC PCM images, the proposed method approached the segmentation performance obtained using the contrast thresholding algorithm, which had been designed specifically for this specific problem. In contrast, trainable segmentation relied solely on the input data and the corresponding ground truths to figure out the best approach for image segmentation. Moreover, it also performed well for co-culture scenarios, a segmentation task

involving the discrimination between two foreground objects with similar visual attribute, something that is unlikely to be achieved robustly using conventional approaches.

In general, Basic Image Features were found to offer the best performance to processing time ratio, even more so in the case of hESC PCM images segmentation. This is not surprising; BIFs are high level descriptors, capable of capturing structures that are not readily encoded in the information provided by either intensity or contrast features. Many features previously described in the literature could potentially further increase the palette of image structures and properties used as input for pixel classification. Those include local binary patterns, Hessian matrix characteristics, histogram of gradients or Gabor filters. For this study, BIFs were chosen over other features due to the ability to encode them in very compact histograms of just seven bins. Given that the time taken to construct local histogram is partially dependent on the number of bins considered (Figure 4-16), compact histograms are usually preferred. Likewise, a major advantage of BIFs over other feature types that can take hundreds of bins to encode is that smaller feature vectors usually lead to shorter training and prediction times as well as lower memory requirements. The ability of BIFs to capture informative image patterns and structures, combined with their compact representation, made them the ideal high level features to investigate in this work.

For the vast majority of the results presented here, using local histograms to encode image features resulted in improved segmentation performance when compared to schemes usually adopted by other trainable segmentation implementations (Sommer et al., 2011a) where a single value per feature and per scale is used for the construction of the pixel feature vector. This agrees with the intuitive perception that image features, and in particular texture, are defined over multiple pixels and that isolated pixels cannot possibly capture that information. Similarly, multi-scale local histograms tended to perform better than single-scale schemes. This suggests that encoding information at various level of granularity was necessary to achieve a proper image understanding.

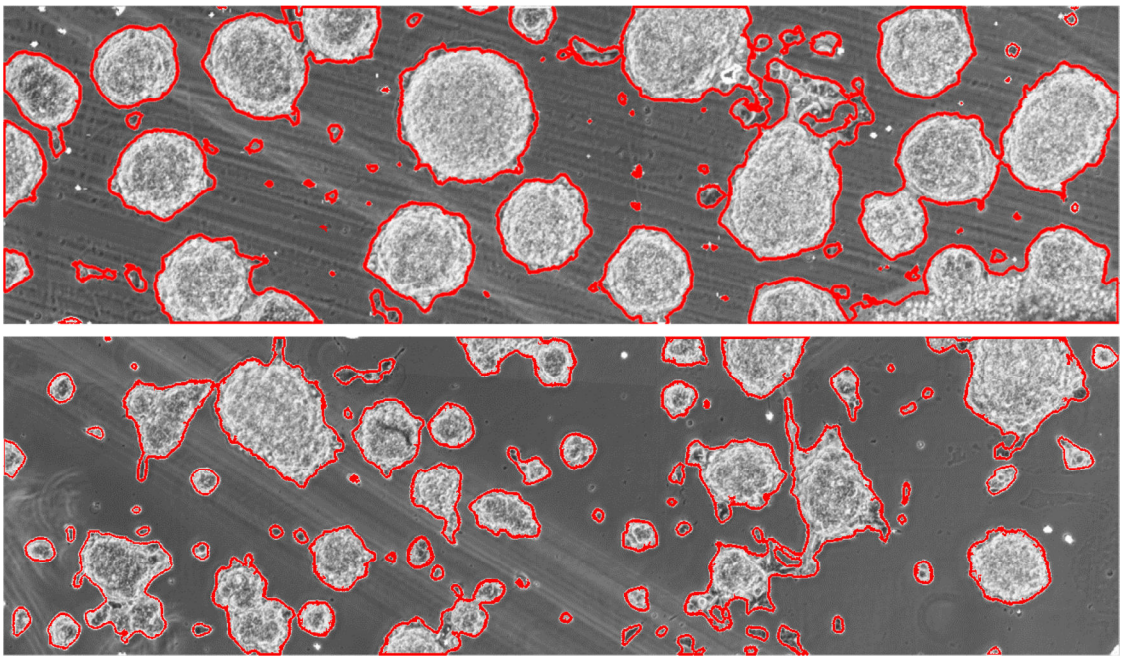


Figure 4-33 Robustness to structured background noise and unusual cell visual features. mESC cultures in the microfabricated culture device described in chapter 5 display a unusual 'stressed' phenotype. Multiple instances of structured background noise are shown, including scratches and depositions. The algorithm trained only on 'normal' examples was able to detect the cells accurately, showing good robustness to these artefacts. Segmentation outline is shown in red.

The trainable segmentation scheme presented in this chapter was also found to be robust to drastic changes in cell phenotype and structured background noise (Figure 4-33). Even though the data used for training only consisted of 'normal' examples, cells with unexpected visual features (e.g. bright, uniform objects) in unseen image were accurately detected. Moreover, structured noise such as scratches, non-cellular depositions, and illumination artefacts were correctly classified as background pixels. These qualitative observations based on the inspection of segmentation results will need to be confirmed by a thorough quantitative characterisation in order to determine to what extent the trainable segmentation scheme can accommodate such artefacts.

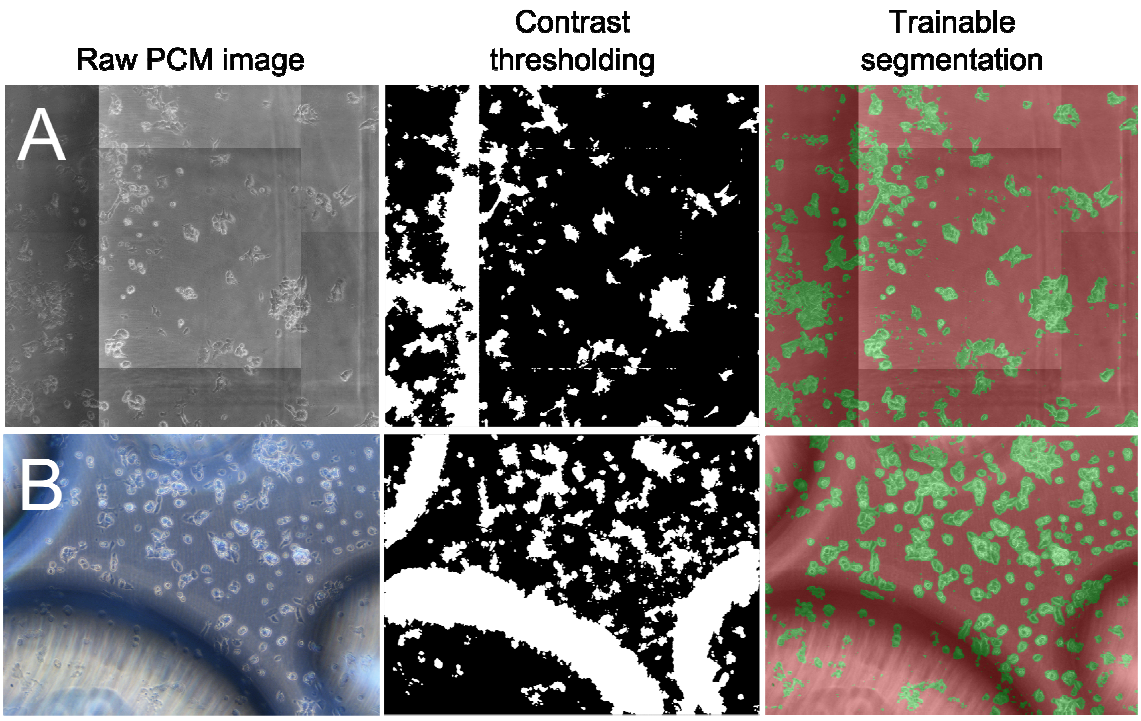


Figure 4-34 Trainable segmentation, given suitable annotations, outperforms contrast thresholding in extreme cases. (A) Illumination artefacts created by stitching multiple images acquired using different exposure times and (B) presence of bubbles. For contrast thresholding, the nseSTDEV with optimal parameter values as determined in section 2.3.2 was used. For trainable segmentation, hard-edged BIFs histograms computed at scales 1+4 and for a window width of 15 were employed. Green and red pixels represent cell and background regions, respectively.

The trainable segmentation approach, when given relevant training annotations, can handle extreme cases of imaging artefacts better than the contrast thresholding algorithm. A first example is an image with illumination artefacts created by the stitching of multiple images that were acquired using different exposure times (Figure 4-34.A). This situation arose when automatic exposure control was enabled in the acquisition software. The sudden changes in illumination intensity resulted in the emergence of artificial edges in the image, which were incorrectly detected by the contrast thresholding algorithm as cell pixels. Conversely, the trainable segmentation method was able to accurately detect the cells without being sensitive to these artificial edges. Another extreme case was the presence of bubbles in the image, a relatively frequent occurrence when culturing cells in a microfabricated device such as the one described in chapter 5 (Figure 4-34.B). The contrast thresholding algorithm detected these bubbles as cells, thus generating a large number of false positives. The trainable segmentation approach was robust to this type of structured noise and, again, was able to correctly detect cellular regions. In order for the trainable segmentation algorithm to achieve such high performance in these extreme cases, it first required the annotation of relevant pixels and the re-training of the classifier. This process only took a few seconds using a bespoke interactive trainable segmentation user graphical interface developed in MATLAB (Figure 4-35), which was inspired by the approach taken by Ilastik (Sommer et al., 2011b).

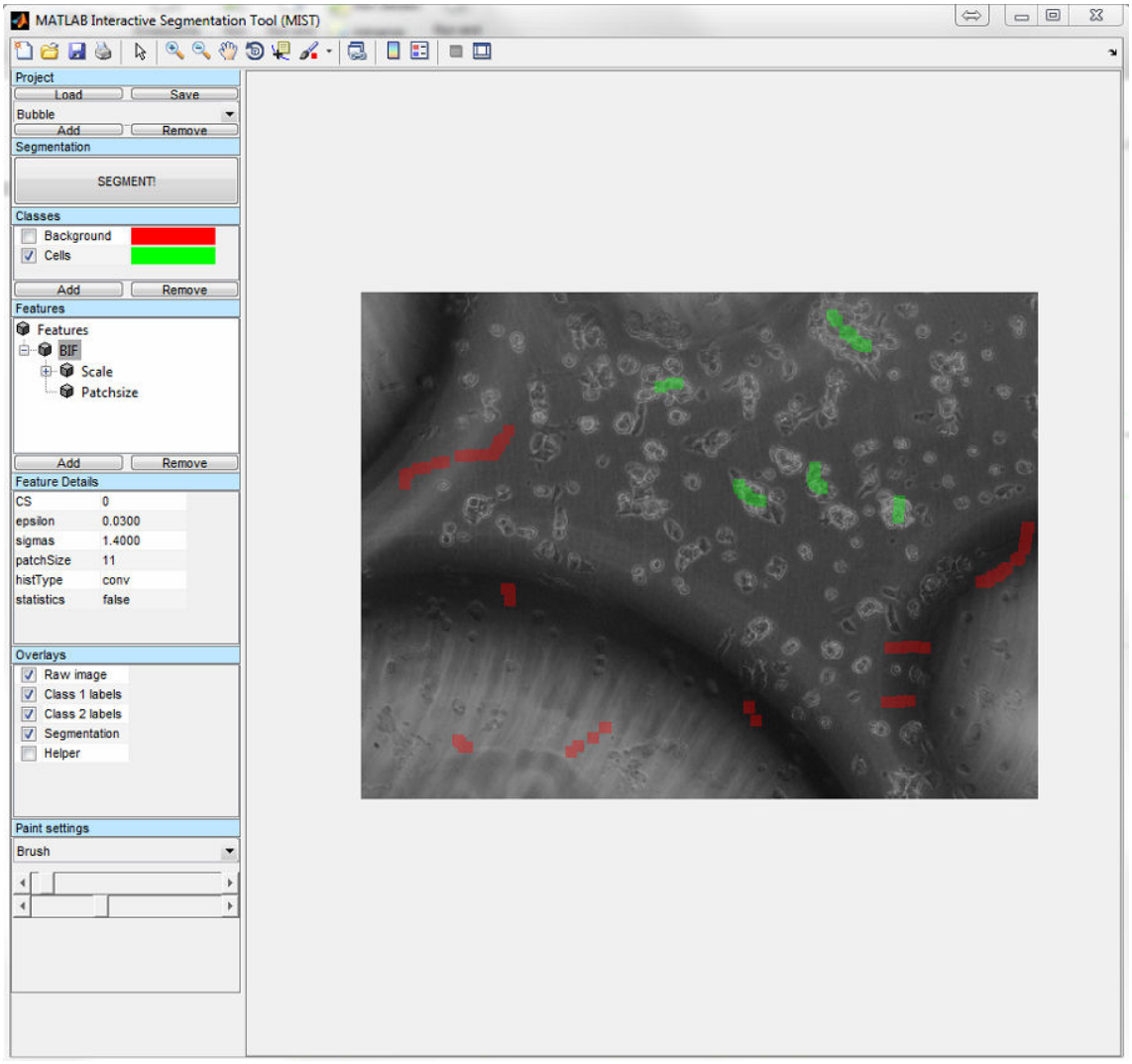


Figure 4-35 Bespoke graphical user interface for interactive trainable segmentation in MATLAB. The active image is the one shown in Figure 4-34.B, with examples of user annotations (red and green for background and cell annotations, respectively).

This study focused on the use of the random forest algorithm for pixel classification. This choice was informed by preliminary experiments, and by evidence in the literature that this classifier can deliver state-of-the-art performance in certain segmentation tasks. Other classifiers such as ADABOOST, artificial neuronal networks, or naïve Bayes could potentially perform as well, or supplant, the random classifier used. It would therefore be interesting to perform a comparison study, which could be used to identify the advantages and weaknesses of various classifiers for pixel-wise classification. Limited preliminary data suggested that random forest performed better than the aforementioned classifiers. However, this was done without any attempt to optimise parameters, which would be needed in order to perform a fair comparison between different classifiers.

Another avenue for the further improvement of segmentation performance could be the use of the recently described context sensitive segmentation. This operation is based on an iterative classification scheme whereby the results of previous iterations are used as inputs (in the form of confidence maps) for future classification iterations. Performance in some segmentations

tasks was significantly improved when using context sensitive segmentation (Seyedhosseini et al., 2011).

All the results presented in this chapter were obtained by randomly sampling 50,000 pixels across the set of images considered. This number was chosen as it represented a reasonable trade-off between the quality of the resulting classifier and the training time. Indeed, training time was proportional to the number of examples (i.e. pixels) considered. Preliminary results had shown that increasing the number of sampled pixels did not significantly increase the segmentation performance. However, below a certain amount of training pixels (usually found to be around 10,000), the segmentation performance rapidly deteriorated. 50,000 pixels represented 1.6% and 1% of the total number of pixels in the mESC and hESC datasets, respectively. Annotating this low number of pixels for training would only take a few minutes at most. Sparse annotations are a major advantage of trainable segmentation over other conventional methods.

Of course, trainable segmentation only makes sense if the training data used can be trusted. In this study, images were manually annotated by human experts. This approach is, in a sense, highly subjective and biased. Indeed, the classifier will be trained to recognise cells in the same way as humans do, including mistakes. However, manual annotations remain the gold standard in light microscopy image analysis due to the lack of viable alternatives. It would in theory be possible to improve the contrast of the cellular object of interest by various means (e.g. fluorescence stains), thus allowing the generation of ground truth data without human intervention. In practice, it is usually not as straightforward. Fluorescence images still need to be binarised before being usable as ground truths. This is usually carried out using a threshold operation, whose output is depending on the choice of an appropriate threshold value. Not only does this introduce a potential bias, but the result is unlikely to be perfect due to the nature of fluorescence microscopy images (e.g. inhomogeneous or non-specific staining, auto-fluorescence). Besides, the additional handling and tampering required to properly label the cells is likely to impact on their phenotype and morphology. The validation of light microscopy image processing techniques remains an on-going issue that will hopefully be tackled and resolved in the near future. Alternatively, performance issues resulting from a small fractions of the ground truth data being mislabelled could be alleviated by using classifiers that are designed to accommodate relatively high noise in the training data, such as RobustBOOST (Freund, 2009).

In summary, the trainable segmentation approach presented in this chapter provided a versatile framework for the segmentation of cell culture PCM images. In chapter 5, its combination with automated image acquisition capabilities to enable monitoring of pluripotent stem cells in a microfabricated device is presented.

Chapter 5

Integration of imaging-based monitoring with a microfabricated device for cell culture process development

The Human embryonic stem cell cultures described in this chapter were carried out by Rhys Macown (Biochemical Engineering, UCL). The mouse embryonic stem cell cultures in the microfabricated bioreactor were carried out by Alexandre Super (Biochemical Engineering, UCL). The processing of the PCM images and subsequent data analysis are the original contributions described in this chapter.

The scaling down of cell culture systems is in principle a very attractive proposition. A smaller scale factor allows carrying out a comparatively larger number of experiments for an equivalent footprint. Smaller reagents and culture medium volume can, in most cases, result in overall costs reduction. Moreover, devices built using microstructures (e.g. sub-millimetre wide channels) can take advantage of phenomena only observed at the micro-scale, such as laminar flow profiles that can be exploited to finely control the biochemical microenvironment of the cells.

However, the design of devices for small scale culture of adherent cells in general, and that of pluripotent stem cells in particular, is highly challenging. Indeed, many of the design principles of the microfluidics field cannot be applied directly, due to the relative fragility of the cells (e.g. shear stress, non-biocompatible materials). As the field matures through the use of new materials or the refinement of designs, many of the major hurdles will be overcome. At any stage of that process, it will be necessary to possess suitable analytical capabilities so that the outcome of experiments can be thoroughly evaluated and eventually compared to the current gold standards at conventional laboratory scales. Yet, for reasons outlined below, it remains difficult to characterise the behaviours of cells cultured in small scale devices.

5.1 Microfabricated device for cell culture process development

The microfabricated culture device was originally designed by Marcel Reichen for the scale-down of embryonic stem cells processes (Reichen et al., 2012). The device was made up of a combination of disposable and potentially re-usable components (Figure 5-1).

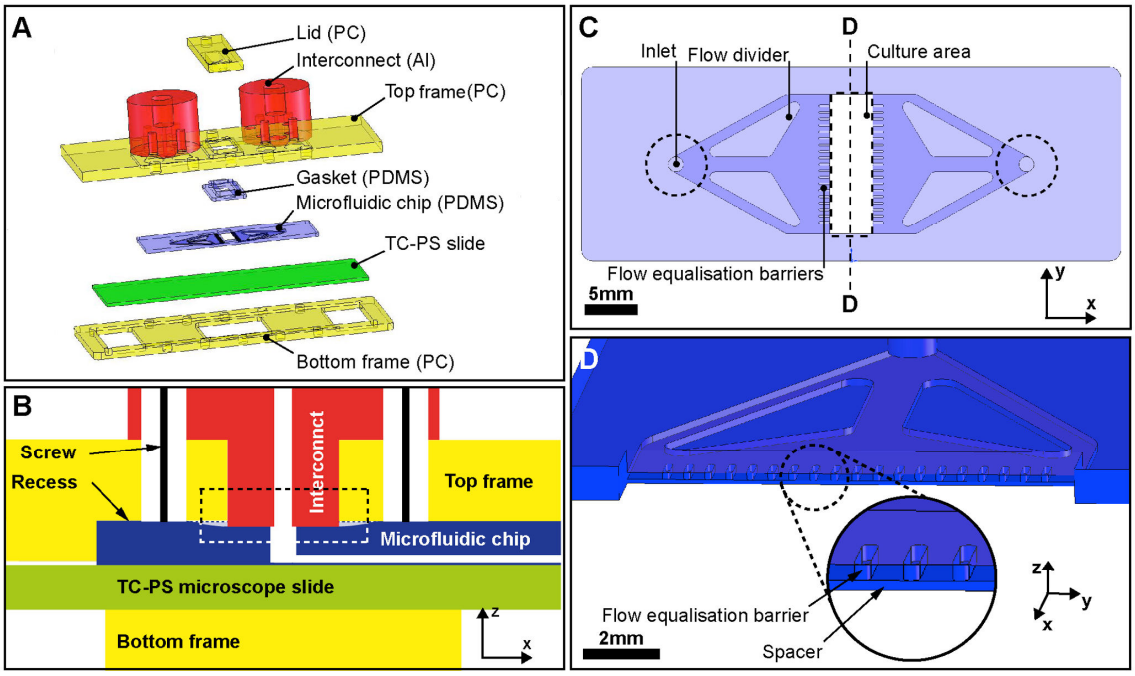


Figure 5-1 The microfabricated culture device. (A) Exploded view illustrating all components of the device. PC is polycarbonate, Al aluminium and PDMS poly(dimethylsiloxane) (B) Longitudinal view illustrating the compression of the PDMS microfluidic chip by the interconnect (dashed rectangle). (C) Top-down view of the microfluidic PDMS chip. Dashed lines represent the footprint of the interconnects and that of the polycarbonate removable lid. (D) Cross-sectional view of the PDMS microfluidic chip. The insert shows a magnified region that illustrates the flow equalisation barrier and the spacer used to elevate the flow in relation to the cells. This figure was reproduced from (Reichen et al., 2012).

At the heart of the device, a poly(dimethylsiloxane) (PDMS) chip defined the contours of the culture chamber and housed channels used to perfuse culture medium. The culture chamber had an effective growth area of 0.52 cm^2, which was in between the areas of a single well of a 96-well plate and that of a 48-well plate. The culture chamber was recessed relative to the fluid flow plane so that potentially detrimental hydrodynamic shear stress could be minimised. The length of the chamber in the flow direction was small compared to its width, in order to have cells across the entire chamber exposed to similar concentrations of nutrients and other small molecules. Flow equalization barriers were positioned upstream and downstream the chamber to ensure homogeneous flow patterns across the culture chamber.

The PDMS chip was positioned on top of a standard microscope slide that served as growth substrate for the cells. This allowed the use of any material for the cell attachment as long as it was available in a standard microscope slide format, including cell culture treated polystyrene used in conventional well plates and flasks. The PDMS chip was compressed against the slide by clamping between rigid polycarbonate frames, forming a reversible seal. The culture

chamber was accessible through a cut-out in the top polycarbonate frame. This feature made the reactor compatible with standard cell culture protocols requiring pipetting (e.g. for immunochemistry). Moreover, it allowed seeding cells directly into the culture chamber (static seeding) instead of relying on dynamic seeding (i.e. cells flown through channels) as typical for microfluidics devices. Dynamic seeding is associated with various issues such as hydrodynamic shear stress and lack of control over the number of cells reaching the culture chamber (Kim et al., 2007). During perfusion, a removable polycarbonate lid defined the height of the chamber (~450 μm) and was used to seal the device by compression against a PDMS gasket. Aluminium (and polycarbonate in an earlier iteration of the design) interconnects were used as an interface with the macro-world, through standard M6 fittings for connection with tubing for perfusion.

To summarise, this design had multiple advantages over previously described microfluidic cell culture devices. Standard seeding, cell recovery or analytical protocols could be carried out in the device due to the removable polycarbonate lid (allowing direct access to the culture chamber). The use of a microscope slide as growth substrate made it possible to use almost any material. Most importantly, it allowed the use of standard and thoroughly validated tissue culture polystyrene whereas a majority of microfluidic devices rely on either PDMS or glass as growth substrates.

5.1.1 The analytical bottleneck and the imaging-based solution

The microfabricated device described above offered a high degree of control over the microenvironment of the cells by allowing the continuous perfusion of culture medium. This should, in principle, make it more suitable for cell culture than conventional vessels (e.g. flasks or well-plates). However, the miniaturisation process comes at a price: conventional analytical methods might not be applicable due to either physical restrictions (e.g. potentiometric electrodes cannot be easily fabricated below a certain size) or to the low number of cells in culture. For indirect measurements of cell activity and function (e.g. pH, dissolved oxygen), this can be partially alleviated by using optical sensors (see section 1.2.7). Cell monitoring is more challenging. Indeed, even methods relying on prior cell detachment (e.g. flow cytometry, PCR) are often unsuitable as the number of cells in the device might be well below their detection limit. This analytical bottleneck poses a fundamental problem when it comes to the usefulness of small scale culture devices for the generation of statistically sound data but also for their comparison with conventional laboratory scale systems.

Phase contrast microscopy (PCM) image processing was shown to be suitable for the monitoring of adherent cell cultures in conventional culture vessels (see chapter 3). This approach was thought to be a perfect fit for the microfabricated device. Because the area of the culture chamber is only 0.52 cm^2, it is possible to image the entire culture chamber frequently.

By doing so, it is no longer necessary to randomly sample the culture but instead the entire population can be analysed, thus drastically improving the precision of the measurements (e.g. confluency).

By its nature, the imaging-based monitoring approach is scale-independent, as long as it is possible to acquire cell images of appropriate quality. However, the proposed PCM image segmentation algorithms also relies on specific image features remaining constant, such as the contrast difference between the cells and the image background. When considering cultures in conventional vessels, this assumption remains true most of the time during the course of a given experiment, but also across multiple experiments following the same protocol. However, this may not necessarily hold true for culture experiments carried out in the microfabricated device, mostly because of their long-term nature. Indeed, a continuous flow perfusion allows to maintain cells in a viable state for periods of time that far exceed what is usually done in conventional systems, with cultures time ranging from 3 days to more than 6 days without detachment. Such prolonged cultures are likely to result in cell visual and morphological features that are not usually observed in conventional culture vessels (examples are shown in Figure 4-2 and Figure 4-33). Between seeding and the end of the culture, these features vary drastically, which prevent the use of the contrast thresholding algorithms presented in chapter 2. The trainable PCM image segmentation algorithm described in chapter 4, is to be employed instead. The machine learning approach offers a high versatility and could accommodate the different phenotypes exhibited by the cells during on-chip cultures, regardless of how distinct they were from one another (see section 4.4). Moreover, trainable segmentation was found not to be sensitive to the presence of defects in images, such as scratches.

In this chapter, the use of a trainable PCM segmentation algorithm for the monitoring of cultures in a previously described microfabricated culture device was investigated. In addition, aspects related to the automation of the culture process, including image acquisition and subsequent analysis, are also discussed.

5.2 Material and methods

5.2.1 Human embryonic stem cell cultures in the microfabricated device

The experiments to generate images of Human embryonic stem cells (hESC) cultured in the microfabricated device were carried out by Rhys Macown (UCL, Biochemical Engineering).

Shef-3 hESC (passage number <70) were cultured on a layer of mouse primary embryonic fibroblasts (MEFs, passage number < 5). MEFs were maintained in the medium formulation described in Table 5-1. Before cultures, the MEFs were inactivated by incubation for 2 hours

after supplementation of the culture medium with 1 mg ml^{-1} of mitomycin C (Sigma-Aldrich, UK).

Table 5-1 Mouse embryonic fibroblast feeder cells maintenance medium formulation

Component	Concentration
DMEM (Invitrogen, UK)	-
Heat inactivated foetal bovine serum (Invitrogen, UK)	10% v/v
Modified Eagle Medium Non-Essential Amino Acids (Invitrogen, UK)	1% v/v

For passaging, hESC cultures were incubated for 5 minutes with a 0.025 mg ml^{-1} collagenase solution (Invitrogen, UK) and then re-placed in regular hESC culture medium (see Table 5-2). The culture was then dissected into colony fragments using a Pasteur pipette.

The microfabricated device was assembled using sterile components (either autoclaved or purchased pre-sterilized) in a biosafety cabinet. In open configuration (i.e. without the polycarbonate lid), approximately 15'000 MEF feeders were seeded into the culture chamber. The device was then transferred into a cell culture incubator (37 °C, 5% CO_2) for incubation overnight so that the cells could properly settle down and attach. After a medium exchange, ~20 dissected hESC colonies were seeded on top of the MEFs layer. Cells were incubated for another 24 hours to enable the proper attachment of the colonies. The device was then sealed using the polycarbonate lid. A syringe pump was connected to the device. Medium was continuously perfused at a flow rate of 300 µl h^{-1} for two days. The total duration of the experiment was therefore three days, one in static culture and two under perfusion.

The cultures were imaged manually each day by transporting the device from the incubator to an inverted microscope (TE2000, Nikon Ltd, UK). The whole culture chamber was imaged at a 10× magnification using a colour camera (Fi-1, Nikon Ltd, UK). Images had a resolution of 1280×960 pixels for a field of view of ~1.2 mm^2.

Table 5-2 Human embryonic stem cell maintenance medium formulation

Component	Concentration
DMEM KnockOut (Invitrogen, UK)	-
Knock out serum replacement (Invitrogen, UK)	15% v/v
Non-essential amino acids (Invitrogen, UK)	1% v/v
L-Glutamine (Invitrogen, UK)	2 mM
β-mercaptoethanol (Sigma-Aldrich, UK)	0.1 mM
FGF2 (4114-TC, R&D Systems, USA)	4 ng ml^{-1}

5.2.2 Mouse embryonic stem cell cultures in the microfabricated device

The experiments to generate images of mouse embryonic stem cells (mESC) cultured in the microfabricated device were carried out by Alexandre Super (UCL, Biochemical engineering).

Prior to the experiment, E14tg2a mouse embryonic stem cells were maintained and cultured as previously described (see section 2.2.1.1). The microfabricated device was assembled using sterile components (either autoclaved or purchased pre-sterilized) in a biosafety cabinet. In open configuration (i.e. without the polycarbonate lid), a solution of 0.1% gelatin was pipetted into the culture chamber and incubated at room temperature for at least 15 minutes. After removal of the gelatin solution, mESC were seeded directly into the chamber at a density of 5×10^5 cell cm^{-2}. The device was transferred to an incubator for 3 hours to let the cells settle and attach properly. The device was then connected to the perfusion system, where flow was driven by varying the head pressure in sealed culture medium bottles using a pressure regulator (ITV0011-2BL-Q, SMC, UK). The device was placed on the stage of an automated phase contrast microscope (Ti-E, Nikon, UK) equipped with a digital colour camera (Fi-1, Nikon, UK) that was used to acquire 10× magnification images at a resolution of 1280×960 pixels, equivalent to a field of view ~1.2 mm^2. An on-stage incubator (Okolab, Italy) was used to control the temperature and maintain it at 37 °C.

The flow rate of the perfusion was set to 300 µl h^{-1}. The cells were cultured for 6 days, during which a LabVIEW routine automatically acquired PCM images of the whole culture chamber (see section 5.2.3 for details).

5.2.3 LabVIEW routine for automated stage control and image acquisition

A LabVIEW (National Instruments, Newbury, UK) virtual instrument (VI) was developed to automate microscope operations (e.g. stage movement, objective changes), fluidics control for perfusion, and image acquisition using the digital camera. The development was carried out in collaboration with Rhys Macown (UCL, Biochemical Engineering).

The microscope module enabled the control of all functions of the automated microscope, including stage position (x-y and z), objective type, fluorescence filter cubes, and illumination intensity (both white light and fluorescence LEDs). The interface with the microscope was done through a set of C++/C# libraries provided by the manufacturer. Similarly, the camera module allowed the display of a continuous video stream, the configuration of the various acquisition settings (e.g. resolution, gain, exposure), and the acquisition of images. Again, C++/C# libraries provided by the manufacturers enabled the communication with the camera via its controller (DS-U2, Nikon, UK). The fluidics was controlled by varying the head pressure of a sealed culture medium bottle. This was achieved by modulating the voltage output of a data acquisition (DAQ) card (National Instruments, US).

5.2.4 Image processing and analysis

5.2.4.1 Segmentation

PCM images were processed using the trainable algorithm described in chapter 4. Hard-edged local basic image features (BIFs) histograms were constructed using the convolution method (see section 4.2.2.3). BIFs were computed at scale (σ) 1, 2, 4, and 8 as described in section 4.2.1.3.

For hESC PCM images, the segmentation task consisted in labelling pixels as one of two classes: hESC colony and other, with the latter class including both image background and mouse embryonic fibroblasts. In the case of mESC PCM images, the pixels were labelled as cell or background. In both cases, the set of images used to train the random forest classifier were from an independent dataset (i.e. not from the monitoring experiment).

5.2.4.2 Image stitching

For the generation of whole-chamber images, individual PCM images had to be stitched together. In the case of the hESC experiments, this was done manually in an image editing software. For mESC cultures, the images were stitched together using a bespoke script. In brief, the position of the stage was recorded for each image acquired. Based on pre-established calibration data, these positions were mapped to a pixel coordinate system, so that the relative position of images was consistent with that of the fields of view during acquisition.

5.2.4.3 hESC colonies characterisation

Whole chamber images were processed using the trainable segmentation algorithm. A connected component analysis was then carried out to detect individual objects (i.e. hESC colonies), which could then be counted. The area of a colony was determined based on the number of pixels of the corresponding object and converted to micrometres using calibration data for the objective, microscope, and camera used.

5.2.4.4 mESC local growth patterns visualisation

PCM images of the culture chamber were first processed individually using the trainable segmentation algorithm. The binary segmentation outputs were then stitched together as described in section 5.2.4.2. A graphical user interface tool allowed the user to place a grid on-top of the image, which defined the size of the local regions to be analysed. Once done, the confluency (i.e. ratio of cell pixel to total number of pixels) for each region was computed. A grayscale heat map where black represented a confluency of 0 and white a confluency of 1 was generated.

For the visualisation of local growth rates, a simple linear regression model was used to describe the evolution of the confluency during the course of the culture for each user-defined region. A heat map was then generated, where each element corresponded to the growth rate (i.e. coefficient of regression).

5.3 Results and discussion

5.3.1 Intermittent monitoring of human embryonic stem cell cultures

Human embryonic stem cells (hESC) cultures in the microfabricated bioreactor were monitored using the trainable segmentation algorithm (see chapter 4) to process PCM images of the culture chamber. The segmentation task was significantly more complex than what was previously achieved using the contrast filter approach (see chapters 2 and 3). Indeed, hESCs are conventionally co-cultured with feeder cells, whose role is to provide factors that are necessary for the maintenance of pluripotency. Feeder cells (in this case inactivated mouse embryonic fibroblasts, or MEFs) and their behaviours are, most of the time, of little relevance to the outcome of an experiment. Instead, hESC colonies are the main objects of interest, and in particular their number and the area they cover. Viable and pluripotent hESC colonies expand as the constituent cells proliferate, thus leading to an increase in measured area. Moreover, it is of interest to track the fate of the colonies as they go through events such as merging, splitting or disappearance (e.g. washed out by the continuous perfusion). The main objective is to characterise hESC growth behaviours in the microfabricated device for comparison with those observed in traditional culture wells.

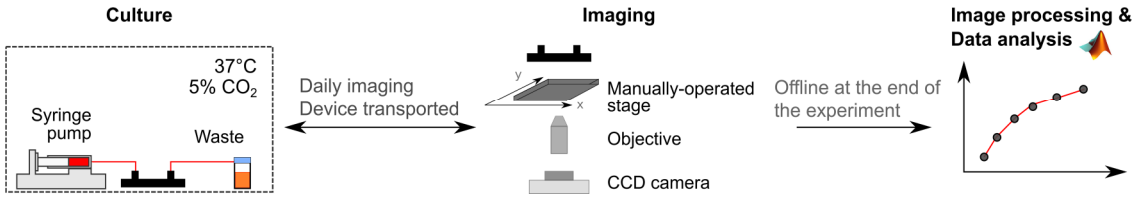


Figure 5-2 Setup for the intermittent monitoring of cultures of Human embryonic stem cells (hESC) in the microfabricated device. The device was kept in a conventional cell culture incubator, together with the syringe pump used to drive the culture medium perfusion. Image was performed by physically transporting the device to a phase contrast microscope where the culture chamber was imaged. Image processing and data analysis were performed offline at the end of the experiment.

For these experiments, an intermittent monitoring strategy was implemented. The culture device was kept in a conventional cell culture incubator, together with the syringe pump used to drive the perfusion flow. For imaging, the culture device was removed from the incubator and transported to an inverted microscope where the culture chamber was imaged. Stage positioning was done manually by allowing enough overlap between individual images so that reconstruction of an image of the whole chamber by stitching would be possible. Image processing and data analysis were performed offline, after the experiment ended.

The appearance and visual features of both mEFs and hESC colonies were comparable between cultures in the microfabricated devices and those carried out in a round-shaped dish (Figure 5-3). The same trainable segmentation algorithm was used to processes images from both vessels.

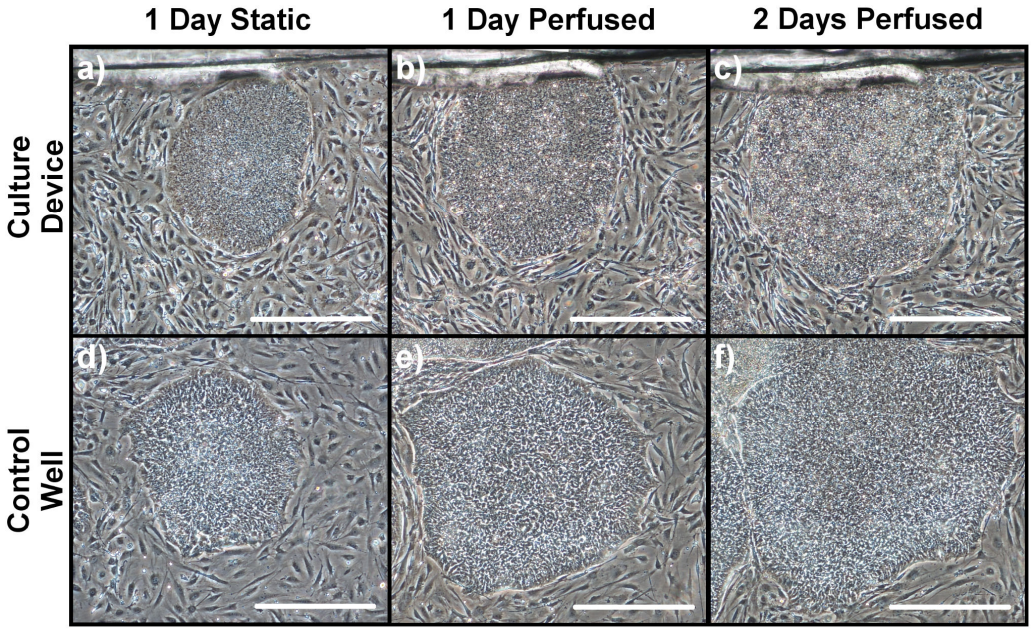


Figure 5-3 Typical images of hESC co-cultured with MEFs in the microfabricated device (a-c) and a conventional culture dish (d-f). In each row, the same colony is shown for three different time points: after 1 day of static culture in open configuration (a,d), after 1 day of perfusion culture in closed configuration (b,e) and after 2 days of perfusion (c,f). Images were acquired at a 4× magnification. The scale bars are 500 μm. Image courtesy of Rhys Macown, adapted from (Reichen et al., 2012).

First iterations of the trainable segmentation algorithm were based on hard-edged local histograms computed using a simple sliding window approach. However, this method was associated with a high computational cost and would take up 50 seconds per image to be processed (see section 4.2.3). Given that the number of images required to cover the whole culture chamber varied from ~20 to ~50 images depending on the magnification used (for 4× and 10× objectives respectively), it could take up to an hour in the worst case to process a single time point. This was deemed acceptable for this experiment, where the cultures were imaged once a day. However such lengthy processing times would not be suitable for online monitoring applications discussed later, whereby cells were imaged once every 30 minutes. Instead, newer iterations of the algorithm use a convolution approach for the construction of local histograms (discussed in details in section 4.2.2.3). While the segmentation output was identical, this approach drastically reduced the processing time per image to less than 5 seconds, thus making it possible to process images of the whole culture chamber within a few minutes only.

Table 5-3 Monitoring of hESC cultures in the microfabricated bioreactor and a conventional cell culture dish (single-well dish with a 2.89 cm^2 culture area)

Culture vessel	Time	Detected colonies [-]	Total colony area [mm^2]	Mean colony area [mm^2]
Microfabricated culture device	Day 1	27	2.9	0.11
	Day 2	25	4.2	0.17
	Day 3	23	6.8	0.30
Conventional culture dish	Day 1	17	2.3	0.13
	Day 2	16	5.0	0.32
	Day 3	10	8.2	0.82

Three metrics for the characterisation of the hESC cultures were computed from the segmentation output: number of detected colonies, the total area covered by the colonies and the mean area of colonies. These metrics allowed performing a direct comparison between a microfabricated device and a conventional single-well culture dish (Table 5-3). The mean area of hESC colonies in both vessels was similar at day 1. As early as day 2, the colonies in the culture dish had a mean area almost double that of colonies in the microfabricated device and a ~3-fold difference could be observed at day 3. These discrepancies could indicate that the culture conditions in the microfabricated bioreactor (e.g. continuous perfusion) were less favourable for hESC colony expansion.

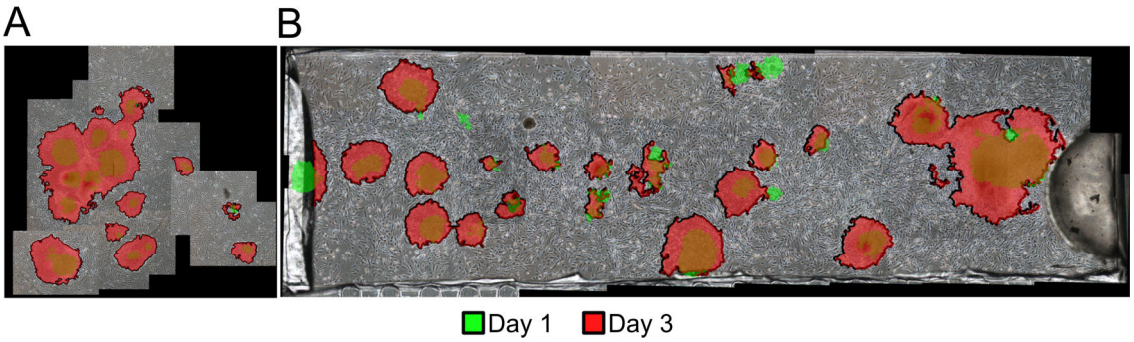


Figure 5-4 Tracking dynamic culture events. Composite images showing detected hESC colonies at day 1 (green) and day 3 (red) overlaid over stitched PCM images of (A) the central area of a conventional culture dish and (B) the entirety of the microfabricated device culture chamber.

The number of colonies in both the single-well dish and the microfabricated culture device decreased as the culture progressed. Such a decrease can be the result of neighbouring colonies merging as they expanded or the disappearance of said colonies following loss of viability and subsequent detachment. To further investigate these phenomena, composite whole-culture images showing the segmentation results for day 1 and day 3 time points together were generated (Figure 5-4). In the case of the culture dish, only the central area of the vessel was imaged as it contained all hESC colonies (Figure 5-4.A). Imaging the rest of the growth area would have prolonged the time cells were left in a non-optimal environment without providing additional data for the analysis. In contrast, the entire culture chamber of the microfabricated device was imaged in a comparatively short amount of time (Figure 5-4.B).

These images showed that in a majority of cases, the decrease in the number of colonies could be attributed to the merging of nearby colonies. However, a few colonies in the microfabricated device disappeared between day 1 and day 3. As no such example could be observed in the culture dish, these disappearances could potentially be attributed to the continuous perfusion washing out loosely attached colonies. A few colonies also appeared to have migrated during that time, while some others became fragmented. This visualisation thus enabled the detection of a variety of events that occur during routine expansion of hES cells. In this case, the colonies were not automatically tracked as imaging was only done on a daily basis, which did not provide the temporal information necessary to accurately track individual objects. Given more frequent imaging, a tool for the automated detection of colonies merging, splitting, migrating, and potentially disappearing could be developed. Nevertheless, this relatively simple approach already highlights the power of imaging-based monitoring towards gaining a better, and more quantitative, understanding of cell culture processes.

The results also showed that trainable segmentation was a good fit for the monitoring of hESC cultures in both a microfabricated culture device and a more conventional culture dish. Despite the relative simplicity of the experimental setup, it was possible to quantitatively measure key culture characteristics and quantitatively assess the performance of the vessels considered. Trainable segmentation appeared to perform reasonably well given a relatively complex segmentation task as all hESC colonies were detected accurately.

5.3.2 Continuous monitoring of mouse embryonic stem cell cultures

Mouse embryonic stem cell cultures were used as a model to demonstrate the feasibility of online (or continuous) monitoring of long-term cultures in the microfabricated device. In order to achieve this goal, the experimental procedure had to be improved. The culture was performed directly on-top of an automated microscope (Figure 5-5.A). An on-stage incubator was employed to maintain optimal growth temperature (37 °C). In the previous intermittent monitoring experiment, a conventional cell culture incubator supplied the necessary oxygen and carbon dioxide mix to ensure optimal growth of the cells. However, controlling the gas atmosphere in the on-stage incubator was not practically feasible. Instead, culture medium saturated in oxygen and carbon dioxide was continuously flown over the cells by modulating the head pressure of a sealed bottle using a cell culture gas mix (5% CO_2 and air) .

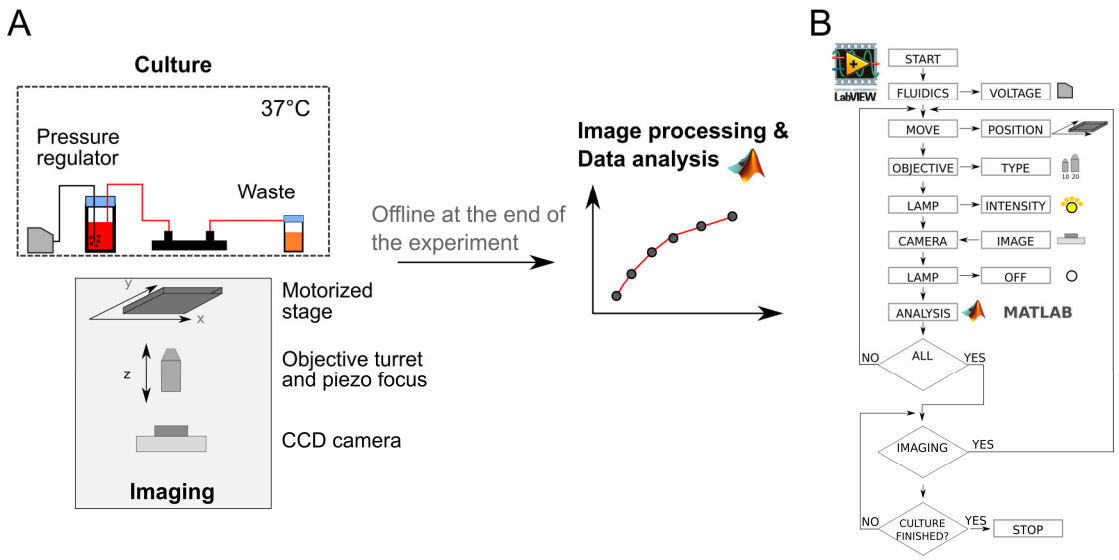


Figure 5-5 Approach for online monitoring of cultures of mouse embryonic stem cells (mESC) in the microfabricated device. (A) Experimental setup. An on-stage reactor was used to maintain an optimal temperature. As the device was placed directly on top the microscope, continuous image acquisition was possible. The various functions of the microscope were fully automated. (B) Flow diagram of the automation routine developed in LabVIEW.

The entire culture setup was automated using a LabVIEW virtual instrument (Figure 5-5.B). The fluidics module was used to set the perfusion flow rate by outputting a voltage to the pressure regulator as based on pre-established calibration data. The main module of the VI was devoted to culture imaging. It was designed so that users could define arbitrary sequences of image acquisition using any combination of settings for each individual microscope component. Possible settings included stage position, the objective type, fluorescence filters and lamp intensity. This functionality was primarily used to setup sequences of image acquisition whereby the location of the motorised stage was varied at each step in order to scan the entire culture chamber. Other uses of this functionality included the imaging of a region of interest at a higher magnification.

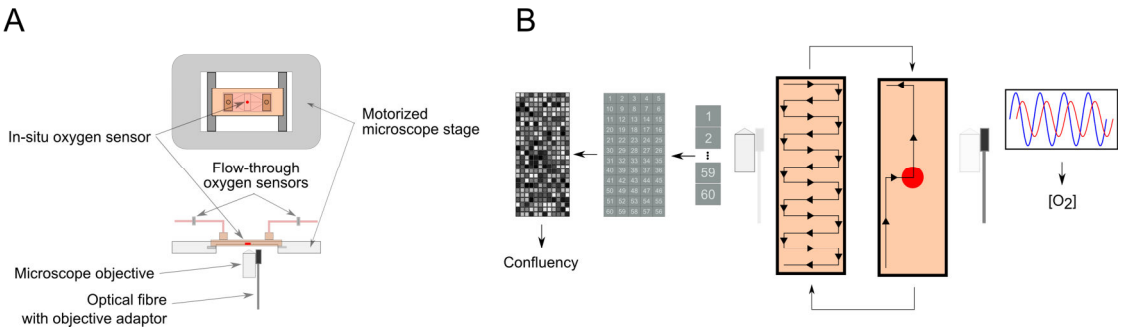


Figure 5-6 Integration of dissolved oxygen monitoring. (A) Schematic of the location for the three sensors used (inlet, outlet and *in-situ*). (B) Imaging of the chamber and positioning of the fibre optic for *in-situ* dissolved oxygen readings.

Even though it was not used for this experiment to avoid excessive CPU load and risk potential crashes, the capability to automatically process the images as they are acquired was implemented. In order to do so, the LabVIEW VI passed the image that was last acquired to a MATLAB scripting node, which processed and analysed the image before returning numerical

outputs (e.g. confluency) to the VI. This information could then be displayed in the graphical interface or stored for offline analysis. Having continuous readings from the culture while the experiment is still running could allow for the early detection of unexpected behaviours.

The experimental setup and its control software also accommodated for dissolved oxygen (DO) monitoring (Figure 5-6.A). A planar optical sensor was placed in the centre of the culture chamber for peri-cellular (i.e. at the cells' level) monitoring of DO. In order to interrogate the sensor and get DO readings, it was necessary to position an optic fibre directly beneath the sensor. This was accomplished by combining a custom-built fibre holder placed directly on the microscope objective (designed by Alexandre Super, Biochemical Engineering UCL) with an automation routine that returned the fibre to the sensor position after each round of chamber imaging (Figure 5-6.B). This was complemented by two flow-cells with optical oxygen sensors placed at the perfusion inlet and outlet. While DO monitoring results are not presented in this study, the imaging-based monitoring approach developed for continuous monitoring in the device had to accommodate the rather large planar sensor patch located in the culture chamber as discussed below.

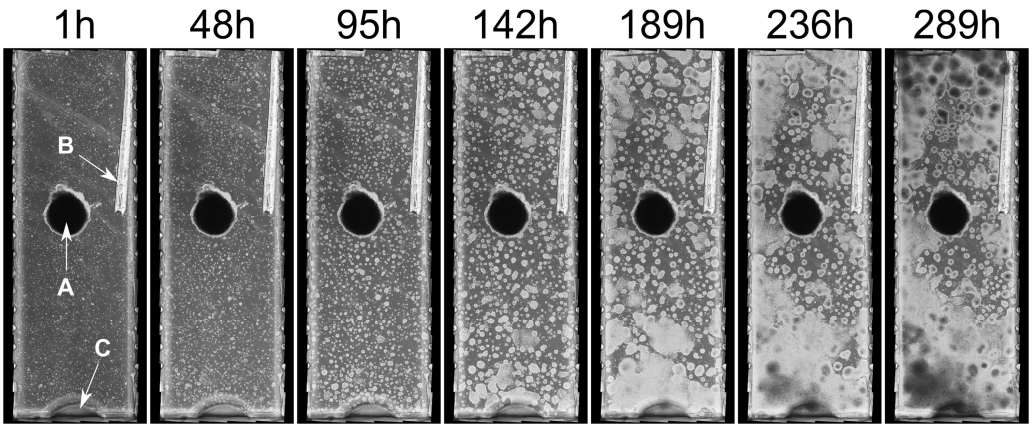


Figure 5-7 Whole culture chamber images at various time points during mESC cultures in the microfabricated culture device. In the first image, A corresponds to the oxygen sensor spot, B to detached PDMS, and C to a indent in the polycarbonate lid.

Images acquired during the course of an experiment were automatically stitched together in MATLAB based on pre-established calibration data that related microscope stage position with image coordinates. This allowed generating full resolution images (~6800×15300 pixels) of the entire culture chamber for each time point considered (Figure 5-7). Besides the cells, these images contained a few structures and features of interest. First, the planar oxygen sensor occupied a large portion of the culture area (Figure 5-7.A). PDMS (the soft polymer used to fabricate the chip housing the culture chamber) could also occasionally peel off and float into the culture chamber (Figure 5-7.B). An indent in the polycarbonate lid also partly obstructed the bottom of the culture chamber (Figure 5-7.C). All of these had to be taken into account at the image processing and analysis stage.

In principle, the segmentation task for PCM images of mESC cultures was significantly less complex when compared to that of hESC PCM images (see section 5.2.1). Indeed, unlike hESC cells that had to be co-cultured with feeder cells, mESCs are usually cultured on their own, thus simplifying the segmentation task to a more conventional foreground versus background scheme. During preliminary development work, the contrast thresholding with *post-hoc* halo removal algorithm (see chapter 2) was evaluated for the segmentation of mESC images acquired in the device. However, two issues that were not usually encountered in conventional culture systems arose: non-cell background artefacts and highly variable cell visual features. Images acquired in the device tended to have much more structured background noise than those acquired from conventional, commercially available culture vessels. Indeed, scratches and other fabrications defects were commonly visible on PCM images of the culture chamber (see Figure 4-1 in chapter 4). This issue was mostly negligible in the hESC culture experiment as the entirety of the culture area was occupied by cells (be it by feeders or hESC colonies). Here, the cells were sparse at the beginning of the culture, thus making these artefacts much more prominent. A simple, naïve approach such as contrast thresholding would likely not be able to discriminate between those background artefacts and cells.

The other issue encountered when processing PCM images of mESC cultured in the device concerned the drastic and seemingly unpredictable changes in cell visual features observed between the various stages of a culture (see Figure 4-2 in chapter 4). While parameter values optimisation could accommodate some of these variations, some phenotypes were so different that they prevented detection of the cells using the contrast thresholding segmentation algorithm as they no longer exhibited the assumed visual features. Approaches using conventional segmentation algorithms were thus unpractical, as they would most likely require the use of separate algorithms and parameter values combinations for the different stages of a culture.

These observations naturally led to the development of the trainable segmentation algorithm described in chapter 4. Indeed, the ability to train a machine learning model to classify pixels based on a variety of local features offered the flexibility and versatility required for this particular segmentation task. The trainable segmentation algorithm was able to accurately label cell pixels on PCM images of mESC cultured in the device regardless of the presence of non-cell background artefacts and despite the unpredictable nature of the cells' visual features (see Figure 4-33 in chapter 4). As such, it was suitable to be used for imaging-based monitoring. However, it was first necessary for the user to determine the regions of the culture chamber that should not be taken into account during image processing and analysis. Indeed as previously noted, images contained structures that were not related to the cells, such as the oxygen sensor spot (Figure 5-7). This was done by developing a bespoke graphical user interface that enabled an easy and quick selection of the regions to be ignored (Figure 5-8.A).

Long-term cultures of mESCs (~5 days) carried out in the microfabricated device were monitored by acquiring images of the entire culture chamber every 30 minutes. Cell proliferation was characterised by determining the culture confluency at each time point considered. Unlike the results showed in chapter 3, there was no error introduced by random sampling in this case as the entire population was considered for each measurement.

The resulting confluency profiles suggested that cells expanded at a significantly lower rate than what was typically observed during expansion experiments in conventional culture vessels (Figure 5-8.B). Inspection of the PCM images showed that cells formed very dense and sparse colonies, which was unusual as mESCs tend to expand quickly and occupy as much growth area as possible. The slow rate of confluency increase, together with the unexpected phenotypes, might have been a response to the continuous perfusion. Indeed, similar cell phenotypes and visual features were previously reported for cultures subjected to continuous or intermittent perfusion (Kim et al., 2006). This response did not appear to correlate with a change in viability or pluripotency of the cells, and thus might be purely morphological. Additional investigations are clearly required to understand the mechanisms underlying this phenomenon.

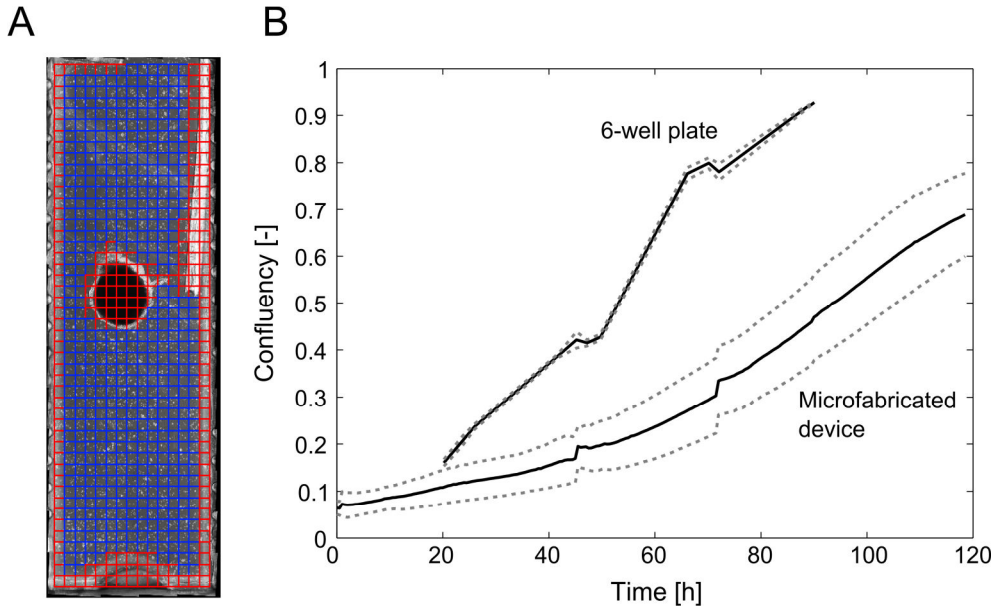


Figure 5-8 On-chip mESC culture confluency monitoring. (A) Software used to specify which tiles to process (blue) or to ignore (red). (B) Confluency profiles in the microfabricated device and in a conventional 6-well culture dish. The dark solid and light dashed lines represent the mean and standard deviation across three independent culture experiments, respectively.

Culture-wide confluency average provided information on the state of the whole cell population. However, it did not provide information on local patterns that might emerge during the course of a culture. Such information was critical for further improvement of the culture device, as it would highlight eventual design or fabrication issues. A simple and straight forward strategy for local growth patterns detection was adopted. The culture chamber was divided in a number of tiles using the same user interface described earlier for the selection of regions to ignore during image processing. For this particular experiments, 47×15 tiles (corresponding to the tiles shown

in Figure 5-8.A) were found to give a good balance between the spatial resolution and the relevance of the information generated. Local confluency was computed for each tile, enabling the construction of a grayscale heat map image where low and high confluency tiles were represented by white and black pixels, respectively. By repeating this process for each time points considered, it was possible to visualise how patterns emerged during cultures in the microfabricated device (Figure 5-9.A). The detection of these patterns was facilitated by incorporating a graphical representation (bar plot) of the average confluency for each row and column.

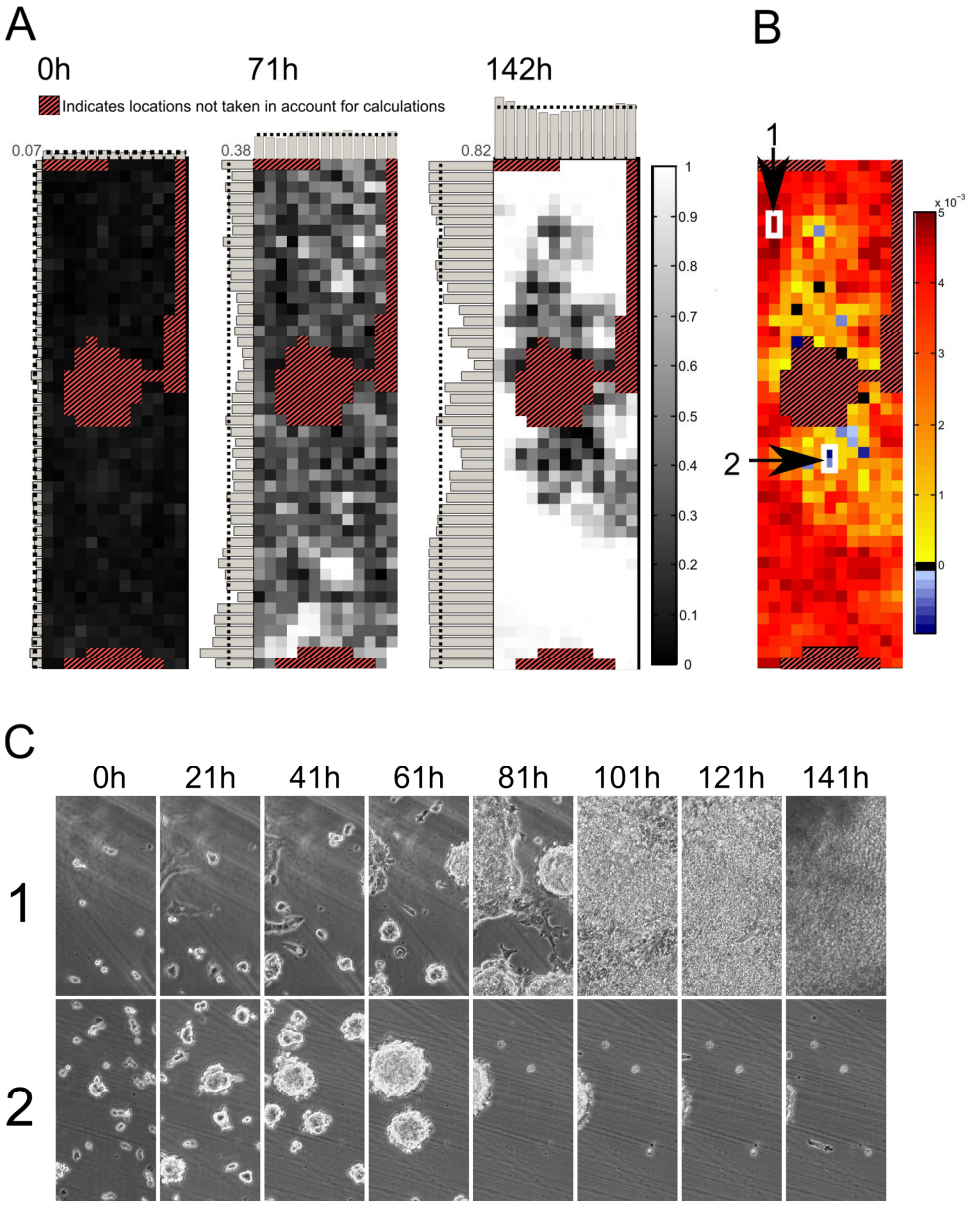


Figure 5-9 Determination of local growth patterns. (A) Heatmap of local confluency measurements for three culture time points (0h, 71h and 142h). The number in the upper left corners of each heatmap denotes the mean confluency for the culture chamber. The bars represent the average confluency for each row and column. The dash line shows the mean confluency for rows and columns. The regions of the heatmap in dashed red were not taken into account for confluency determination. (B) Local confluency increase rates for the entire culture duration. Blue denotes negative increase rates, yellow null growth rates and red high growth rates. (C) Time lapse PCM images of the regions pointed to in (B).

This visualisation tool was used to investigate local growth patterns at three time points (just after seeding, 71 hours, and 142 hours into the culture). Shortly after seeding, the confluency appeared to be uniform across the whole culture chamber, indicating that the protocol used (cells pipetted directly into the culture chamber) was appropriate. As soon as 71 hours after the start of the experiment, patterns emerged as small pockets of high confluency were observed mostly in the top and bottom thirds of the culture chamber. After 142 hours (~5 days) of culture, a large fraction of the culture chamber was nearing 100% confluency. However, the confluency measurements for some of the tiles were comparatively very low, with a few cases close to zero. The regions in question appeared to be clustered around the centre of the chamber, in close proximity to the oxygen-sensing spot.

The local confluency visualisations allowed the user to perform semi-quantitative analysis of cultures in the microfabricated culture device. However, it still required to study multiple images to gain an insight into how patterns formed overtime. To further improve data analysis capabilities, local confluency increase rates were estimated in each tiles by fitting a linear regression model of the confluency as a function of the culture time. This was visualised by the way of a heat map whose pixel colour values were dictated by the slope coefficient of the corresponding regression model (Figure 5-9.B). The regions noted previously as having comparatively low confluency at the end of the culture were also associated with a low confluency increase rate in this new visualisation. In fact, the increase rate for a few of these tiles was negative, effectively showing that cells were at some point present in these regions but subsequently disappeared (Figure 5-9.B, arrow 2). In general, this could indicate that cells either lost viability and detached, or migrated away from these regions. To help determine which of these explanations applied to this case, the user could retrieve time-lapse PCM movies corresponding to these locations by selecting them in the interface. In a region with high confluency increase rate (Figure 5-9.B, arrow 1), the cells first settled down and spread out before entering a proliferation phase around 61 hours into the culture during which cells formed colonies, and eventually reached full coverage by 101 hours (Figure 5-9.C.1). In contrast, the behaviours of the cells in the region with a negative confluency increase rate was significantly different. Small and compact colonies were already forming 21 hours into the culture and increased in size mostly through merging with nearby colonies until 61 hours after seeding, at which point the colonies migrated away, only leaving a few debris in the region (Figure 5-9.C.2). The visual appearance of the colonies suggested that the cells did not flatten and spread as they did in the region with high proliferation, but instead formed a 3-dimensional structure that appeared to be loosely attached to the growth substrate.

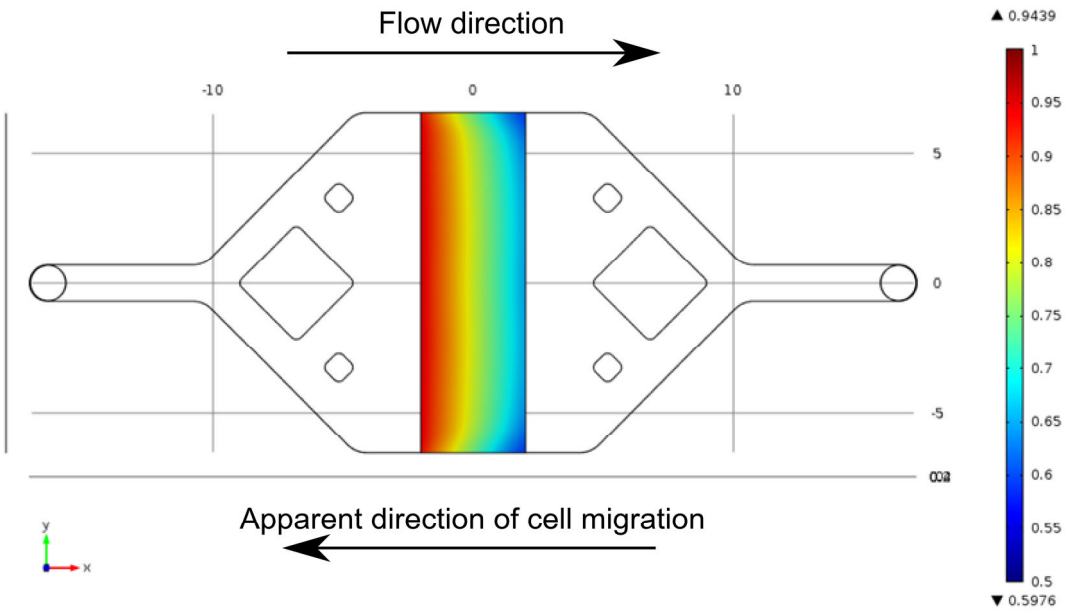


Figure 5-10 Finite element model of oxygen gradient in the microfabricated device. Cell consumption was assumed to be 10^{-7} mol m^{-2} s^{-1}. The numbers shown represent the dissolved oxygen tension (%). Model and image are courtesy of Rhys Macown (UCL, Biochemical Engineering).

More investigations would be required to determine the exact cause of this sudden cell migration in some regions of the culture chamber. A possible explanation could be related to how nutrients, and more specifically oxygen, were supplied to cells cultured in the microfabricated device. In a conventional culture vessel, oxygen usually diffuses from the gas phase to the liquid phase, where it will eventually reach the cells. This process can in most cases be considered homogeneous across the entire culture area. In the culture device, however, the oxygen was supplied by saturation of the perfusion medium. As the culture medium travelled across the culture chamber, its oxygen content would be depleted due to consumption by cells, potentially resulting in an oxygen concentration gradient across the chamber. This was illustrated by a finite element model produced by Rhys Macown (UCL, Biochemical Engineering). The dissolved oxygen tension (DOT) was indeed found to decrease as a function of the distance from the perfusion inlet (Figure 5-10). It could therefore be possible that the cells migrated towards a region of higher oxygen concentration using a chemotactic mechanism. Another possible explanation could involve growth substrate defects that prevented proper cell attachment and spreading (e.g. uneven gelatine coating, mechanical properties altered during assembly or integration of the oxygen sensor patch). Further investigations are required to understand this.

These results demonstrated how an automated image acquisition system combined with a trainable segmentation algorithm could be used for the monitoring of mESC cultures in a microfabricated device. Visualisation tools proved efficient for the determination of local growth patterns and the investigation of events of interest. While the culture information provided by such visualisation needs further analysis, it is an essential first step towards the quantification of cell culture behaviours in microfabricated culture systems.

5.4 Conclusion and outlook

Imaging-based monitoring was shown to be a viable analytical approach for mouse and Human embryonic stem cells cultures in the microfabricated device considered. Even when only intermittent imaging was used, relevant and useful information could be generated from phase contrast microscopy images. Trainable segmentation performed well in cases where conventional approaches, such as the contrast thresholding algorithm described in chapter 2, would not be suitable. Maybe more importantly, processing time was within reason with less than 5 seconds required to process a full resolution image. Imaging-based monitoring was in a sense better suited to the scale of the device, as it was possible to study the entire population of cells instead of relying on random sampling like it was the case for conventional culture vessels.

As such, the methods described in this chapter represent a first step towards the mitigation of the analytical bottleneck that is associated with the miniaturisation of culture systems. Indeed, using a simple metric such as confluency, the performance of the microfabricated device was compared with that of conventional culture vessels thanks to the scale-independent nature of the approach. Likewise, the analysis of local confluency patterns was used as a basis for the detection of cell responses that might have been missed otherwise. In situations where even end-point analyses were challenging to carry out, imaging-based monitoring provided simple yet relevant and useful culture data.

Many aspects of the approach described in this chapter could be further improved to provide more sophisticated culture data. One such improvement could be automated cell tracking across time-lapse microscopy frames. The results described in this work were obtained by essentially taking a fresh snapshot of the culture at different time points, without correlation to previous time points. Instead, individual cellular objects could be tracked during the course of a culture. This would allow, for example, to quantify events such as colony merging, splitting, migration or disappearance. However, tracking such a large number of objects could be computationally very expensive. In addition, it remains challenging to track objects in dynamic environments (e.g. changes in illumination, structured background noise, cell proliferation). Another improvement would be the integration of fluorescence microscopy. The control software could accommodate the acquisition of fluorescence images in addition to that of PCM images. It could therefore be possible to combine PCM and fluorescence microscopy to monitor additional culture characteristics such as pluripotency or degree of differentiation, similar to what was shown in section 3.2.4 for conventional culture vessels.

In this study, image processing and analysis was mostly done offline, after an experiment was concluded. However, the LabVIEW virtual instrument used for automation also accommodated online image processing while a culture experiment was still on-going. This feature was important as it meant that both control functions (e.g. perfusion flow rate, temperature) and

monitoring capabilities were housed in a single program, thus allowing them to interact. A possible application would be the development of culture protocols where decision-making would be automated, based on the current state of a culture as determined by imaging-based monitoring. This is in contrast with most experimental protocols where the sequence of operations is usually based on arbitrary time points, regardless of the state of a culture. While reasonable for simple experiments, an approach relying on arbitrary time schedules breaks down for complex applications such as directed differentiation of pluripotent stem cells where culture-to-culture variations render the different time points effectively meaningless. By combining control and monitoring, experiments could be carried out automatically where the different steps would be based on quantitative measurements such as culture confluency.

Microfabricated devices are a perfect fit for cell biology, offering a degree of control on the environment of the cells that is out of reach of conventional cell culture vessels. The integration of such devices with imaging-based monitoring as described herein could lead to the development of high content platforms that would allow quantitative studies of intricate biological processes. Such advances are desperately needed as lack of suitable tools could impede progress in complex fields such as regenerative medicine.

Chapter 6

Conclusion and recommendations

6.1 Two approaches to PCM images segmentation

Two fundamentally different approaches to PCM image segmentation were described and thoroughly characterised (Figure 6-1). The first approach, contrast thresholding, was developed based on our knowledge of PCM images properties and structures. Each step was purposely designed to handle specific aspects of these images; the contrast filter operation allowed discriminating between background and foreground pixels based on local variance despite similar intensity values while the *post-hoc* dynamic programming step corrected for halo artefacts. In contrast, the problem was approached the other way around with trainable segmentation where a machine learning classifier was trained to classify pixels as either background or foreground based on generic image features. *A-priori* information about PCM images was not explicitly built-in the algorithm, but rather learned implicitly by the classifier.

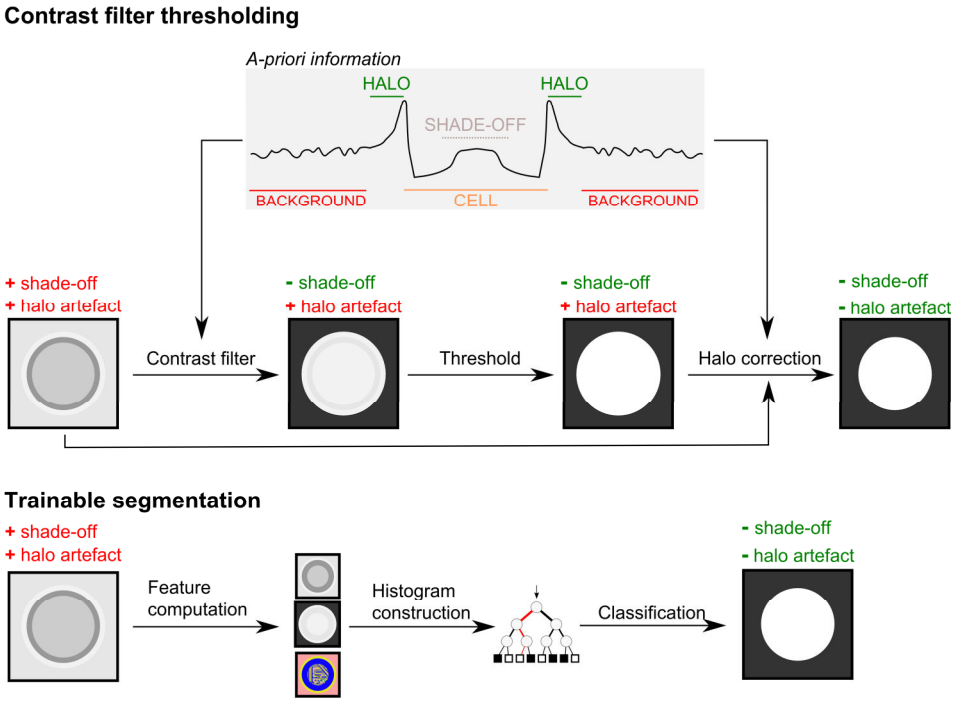


Figure 6-1 Two fundamentally different approaches to PCM image segmentation. Contrast thresholding was devised based on *a-priori* information about PCM images. The contrast filtering and halo correction steps were specifically designed to handle the shade-off and halo artefacts, respectively. Trainable segmentation is not based on any previous knowledge. Instead, a random forest classifier learns how to classify pixels based on generic image features. Both approaches produced comparable end-point results.

The respective strengths and weaknesses of the two approaches indicate that their ideal use cases differ markedly (Table 6-1). Contrast thresholding is highly optimised for applications involving the detection of a single cell type whose visual features conform to the assumptions made concerning PCM images. Moreover, it is most appropriate when a low processing time matters, for example for decision-making during an experiment. It is expected that contrast thresholding would be a suitable choice for most routine cell culture experiments. For more complex applications, trainable segmentation is often preferable despite the significantly longer processing times. Such applications include co-cultures, experiments during which cell visual features are expected to vary drastically, and cultures where structured background noise such as scratches is common occurrence.

Table 6-1 Comparison of contrast thresholding and trainable segmentation. Qualitative assessment of various aspects of segmentation. Qualitative ranking, ranging from + (worst) to + + + (best).

	Contrast thresholding	Trainable segmentation
Processing time[1]	< 1s	~ 4s
Offline phase[2]	+	+ + +
Performance for foreground vs background	+ + +	+ +
Performance for cell type 1 vs cell type 2	N/A	+ + +
Robustness to variations in imaging conditions[3]	+ + +	+ + +
Robustness to structured background noise[4]	+	+ + +
Robustness to variations from expected cell features	+	+ + +

1. For a 1MP image (1280×980 pixels). Representative processing time using a single thread on a modern CPU.
2. Corresponds to parameter tweaking for contrast thresholding and interactive training of the classifier for trainable segmentation
3. Includes different microscope and camera models, changes in illumination, and focusing
4. Includes scratches, non-cellular deposition, and illumination patterns

A major differentiator between the two methods was the offline phase, which corresponded to parameter values optimisation and classifier training for contrast thresholding and trainable segmentation, respectively. By nature, machine learning classifier training only requires sparse inputs, i.e. not all pixels of an image have to be manually annotated. In addition, this training phase was very quick (a few minutes, usually less than one), albeit memory intensive. As such, trainable segmentation could potentially be used for interactive segmentation, whereby the classifier is automatically trained as more data is annotated by the user. A preliminary version of a graphical user interface for interactive segmentation was developed (Figure 4-35). While parameter value optimisation for the contrast thresholding algorithm could also be done based on sparse inputs when using the specialised tool described in this study, the process was still significantly lengthier than classifier training and thus still represented a significant bottleneck. Given that parameter values were likely to be optimised for significantly different applications, the versatility and simple addition of new training examples might give the advantage to trainable segmentation in such situations.

The contrast thresholding algorithm compared favourably with two other methods also based on contrast filtering that were previously described in the literature (Juneau et al., 2013; Topman et al., 2011). The *post-hoc* halo correction step was found to be the main driver for the high performance observed, and its addition to the other two methods resulted in significant increase in their performance. The focus was put on the development of a relatively simple method that would deliver suitable performance while retaining short processing times. It is highly likely that methods employing more advanced techniques such as active contours (Ambühl et al., 2012) would be able to achieve better performance. Indeed, these methods are capable of detecting minute changes in cell contours that would most likely elude simpler approaches such as the proposed contrast thresholding algorithm. However, this is often achieved at the cost of significantly longer processing times. It could thus be argued that both types of approaches are equally valuable and are, in fact, complementary.

Similarly, trainable segmentation and conventional approaches such as contrast thresholding are likely to co-exist for the foreseeable future, as the latter still produce superior results in most cases. However, the ever increasing computational power and the development of more efficient machine learning classifiers might make it possible for these generic approaches to attain performance similar to that of specialised algorithms. Indeed, brute force methods, such as deep learning, already deliver state-of-the-art performance in the field of image classification despite a lack of theoretical grounding and understanding (Cireşan et al., 2012). While image segmentation based on machine learning techniques is a fairly new idea, software packages such as Ilastik and the Weka trainable segmentation plugin for ImageJ (Hall et al., 2009; Sommer et al., 2011b) already represent viable, and often more accessible, alternatives to traditional machine vision and image processing algorithms. In this work, trainable segmentation performance was further improved by the introduction of a multi-scale local histogram scheme for feature encoding, which allowed taking advantage of spatial context information for pixel classification. The increase in performance was most noticeable for complex segmentation scenarios such as the discrimination between two co-cultured cell types.

For relatively simple segmentation tasks (e.g. cells versus background), the performance of the trainable segmentation scheme described here was comparable, if not slightly better, than that of Ilastik and Weka trainable segmentation. As such, a potential user could use any implementation and obtain similar results. However, when processing images that are more complex in nature, such as the co-cultured hESC images presented in this work, it was shown that the use of the novel multi-scale local feature histogram scheme greatly improved segmentation performance.

Going forward, further improvements of either method will be dictated by the emergence of new use cases in cell culture laboratories. One such improvement might be the segmentation of

individual cells within images. The decision to focus on the classification of pixels as 'cell' or 'background' in the whole image without attempting to separate individual cellular objects was made at the beginning of the project. After evaluation of preliminary PCM images of the main cell lines considered in this work (mouse and Human embryonic stem cells), it was clear that segmentation of cells within colonies would be extremely challenging. At the time, global image segmentation was considered satisfactory for the purpose of the project, which was to demonstrate imaging-based monitoring in a wide variety of relevant experimental scenarios. Approaches have been recently proposed for individual cell segmentation in the case where cells remain as relatively well-separated entities with clear boundaries (Su et al., 2013a; Su et al., 2013b). The generalisation of these methods to images of colony-forming cell lines would represent a major milestone and open new avenues for adherent cell culture monitoring.

It is important to specify that the two methods presented in this work are not necessarily suitable for applications outside the scope of those presented. For example, they are not recommended for the segmentation of images acquired using microscopy methods other than PCM (e.g. brightfield or fluorescence microscopy). Likewise, both contrast thresholding and trainable segmentation were devised to handle images of relatively low magnification and other tools should be used to handle high magnification and/or resolution images of single cells (Ambühl et al., 2012) as well as application where sub-pixel resolution is required.

6.2 Towards a standard monitoring approach for adherent cultures

Measurements of relevant cell culture characteristics were carried out based on the output of the contrast thresholding and trainable segmentation algorithms. Confluency, a metric familiar to all cell culture researchers, was an obvious candidate for imaging-based monitoring. The robustness and consistency of the segmentation algorithms allowed using automated confluency measurements for the monitoring of cell proliferation, growth arrest, cell death, and transient morphological changes. Monitoring was achieved in conventional cell culture vessels (6-well plates) as well as in a microfabricated culture device whose culture area was over 18 times smaller. Despite their simplicity, automated confluency measurements allowed to gain valuable insight into the processes studied. In particular, the ability to carry out time-course studies enabled the investigation of dynamic cellular behaviours that would elude end-point analyses (e.g. changes in colony area in response to stress).

Cell density, the gold standard for the characterisation of an adherent cell culture, could not be computed directly from the algorithms' output as only cellular objects were segmented and not individual cells. However, preliminary work showed that the number of cells could be estimated from the confluency measurements if those were corrected for changes in specific cell area as determined by the distance between nuclei-like texture features. Despite the fact that this

approach relied on pre-established calibration data, promising preliminary results suggested that it could be used to determine cell density of colony-forming cell lines solely based on light microscopy images. In the current implementation, the correction consists in dividing the confluency by the mean distance between nuclei-like features. The calibration between cell density and PCC is then obtained by linear regression. An alternative approach would be to perform a multiple regression on the different variables (e.g. confluency, distance between nuclei-like features, and possibly others) using a machine learning algorithm such as random forest. Training would require a substantial amount of data but the resulting model is likely to be more accurate and to generalise better than the current approach.

The segmentation output was also used to perform simple morphometric analyses, i.e. characterisation of the shape of cellular objects. While it provided useful information, morphometric analysis was limited to images from early cultures stages where cellular objects were well-separated. As the cultures progressed, colonies tended to merge and formed objects that were larger than the dimensions of the field of view, thus making their characterisation impossible based on single images. The approach could potentially be improved by considering multiple fields of view during morphometric analysis.

PCM segmentation was also used to facilitate the interpretation and analysis of fluorescence microscopy images. The main idea behind this application was that it is not always possible to have a fluorescence channel devoted to a marker used for the detection of the boundaries of cellular objects. By acquiring fluorescence and PCM images of the same field of view, the segmentation output for the latter could be used to determine which regions of the former corresponded to cells objects that were not expressing a fluorescent marker. Rudimentary analysis of gene expression patterns in response to changes in culture condition was successfully demonstrated.

The examples of applications discussed here showed the vast amount of culture information provided by imaging-based monitoring. Furthermore, these experiments could have been carried out in any research laboratory with the only requirements being a phase contrast microscope, a digital camera for image acquisition, and possibly the equipment necessary for fluorescence microscopy. Likewise, imaging-based monitoring was compatible with standard cell culture ware of any scale so long as it could be adequately imaged using phase contrast microscopy. The only changes in experimental workflows involved the acquisition of images during visual inspection of the culture, which had to be carried out anyway. Imaging-based monitoring thus allowed quantitatively documenting every step of cell culture activity or process. This will provide in the future a sound basis for decision-making and the evaluation of experimental outcome. Moreover, it offers a standardised framework for the communication of cell culture data and could thus improve the robustness and reproducibility of experimental protocols.

The establishment of standard approaches for adherent culture monitoring will rely on open, freely available software and algorithms. Packages such as ImageJ already demonstrated how the scientific and industrial communities can come together to develop very powerful tools. Inter-operability and platform-independence are also necessary to avoid a proliferation of similar but incompatible tools and to enable as many researchers as possible to access and use said tools. The software for this work was developed primarily in MATLAB, a proprietary scientific computing language, as it allowed rapid prototyping of the concept. Ultimately, the entire codebase should be ported to other, non-proprietary platforms. The contrast thresholding and halo correction algorithms were already ported to Java, so it can be used directly as an ImageJ plugin. The source-code for all the components related to the contrast thresholding approach was released under a permissive open-source license, with that of the trainable segmentation algorithm to follow soon. Distribution of the code is critical as it enables other researchers to scrutinise, validate, and improve it.

6.3 The next step in adherent cell bioprocessing

Adherent cell culture has conventionally been a manual and thus very laborious endeavour. The many steps involved in any experimental protocol have to be carried out by thoroughly trained operators. As a consequence, throughput is generally very low. Moreover, it also introduces variability and reliability issues. These issues are exacerbated as the complexity of the process increases. Indeed, the progress in fields such as regenerative medicine or tissue engineering, which rely on the fine control of culture conditions, is impeded by the lack of suitable tools and approaches. One possible approach to this problem is the automation of existing cell culture protocols and workflows. Commercial vendors such as TAP biosystems[6] offer a variety of product that completely automate routine operations such as medium exchange or cell passaging. These systems would also be a good fit for the monitoring approaches described in this project as they can optionally be equipped with light microscopy capabilities. While this automation approach somewhat alleviates the reliability and throughput issue, it does not address the need for advancements in how cell culture is performed at a more fundamental level.

There are not too many opportunities to drastically alter culture vessels at a conventional scale. In contrast, miniaturisation might allow to completely re-think how adherent cell culture is performed. Indeed, downscaling to the micro-scale comes with many advantages, with the most relevant to cell culture being the ability to have liquid flows that are well-understood and characterised, due their highly laminar nature. Applications include the supply of controlled concentrations of nutrients and small molecules to the cells or the generation of chemical gradients across the culture area. Such capabilities are very attractive and they might be used to

[6] http://www.tapbiosystems.com/ (last accessed 01.07.2014)

better mimic the complex and highly dynamic niches found *in-vivo*. However, miniaturisation is often at the expense of the information that can be readily obtained from the process. Indeed, the small volumes and the number of cells often preclude the use of conventional end-point analytical methods.

Imaging-based monitoring approaches are thus of particular interest. Micro-scale culture devices can often be fabricated using optically transparent material, allowing the cells to be imaged using light microscopy methods as they would be in conventional culture vessels. Based on these images, culture characteristics can then be determined using image processing and analysis. This approach was taken to monitor the cells in a previously reported novel microfabricated cell culture device. For this application, the trainable segmentation algorithm was preferred to the contrast thresholding approach due to the nature of the cultures considered (e.g. hESC co-cultured with mouse embryonic fibroblasts) as well as the highly variable cell phenotypes resulting from long-term culture under continuous perfusion (possibly due to prolonged exposure to flow-induced shear stress). This allowed the generation of quantitative culture information for the evaluation of the device performance in comparison to what is typically achieved using conventional culture vessels. When combined with automated image acquisition, this approach allowed the continuous monitoring of the entire culture chamber. This enabled the determination of culture confluency without sampling, thus increasing the reliability of the measurement. Moreover, it was possible to determine local changes in confluency, which could in turn be used to identify the emergence of local growth patterns. This information could be of high value for the further improvement of the device and to gain a better understanding of cell response to dynamic culture environments. The measurements carried out for cultures in the microfabricated device could be significantly improved by tracking objects across the entire experiment, thus informing on key events such as colony merging or splitting.

It is clear that the next evolution should be one where imaging-based monitoring is fully integrated with all other control systems of a microfabricated culture device. Culture information, acquired online through image processing and analysis, would be used as the basis for the control of various culture parameters such as perfusion flow rate, medium composition, gas atmosphere or even temperature. This would allow devising fully automated protocols where decision-making is partly or entirely based on culture characteristics. Such a system could potentially help alleviate the reproducibility issues for the more complex protocols (e.g. differentiation of pluripotent stem cells) as variations between cultures would not impact on the experiment because the protocol would be dynamically on the fly to take them into account. In addition, the automation of decision-making would also minimise variations introduced by qualitative observations by the operator. Likewise, it would greatly facilitate the screening of a large number of conditions. The work presented in this document represents an essential and significant step towards the realisation of such an ideal bioprocessing platform.

6.4 Summary of main contributions, novelty, and advances

The work presented in this book is multidisciplinary in nature, covering fields such as cell biology, biotechnology, image processing, and machine learning. This final section summarises the main contributions of this work and further clarifies some of the design choices made for the image processing algorithms.

The individual components of the contrast thresholding method were based on well-understood principles in image processing and are thus not novel in themselves. Contrast filters were employed to highlight regions of high intensity variations, previously shown to correspond to cell-containing regions. The contrast image was thresholded to obtain a rough segmentation mask, which was then corrected for halo artefacts using a gradient-steered contour shrinking algorithm. Although similar in principle to active contour methods, the proposed algorithm is several orders of magnitudes faster than said methods due to its low complexity. This was made possible by sacrificing generalisation of the underlying algorithms when possible due to *a-priori* knowledge of the structure and properties of PCM images. For example, the image gradient was estimated by convolution with edge-sensitive Kirsch kernels, a method that is sub-optimal in many scenarios but whose results were found to be suitable for this particular problem, with the added benefits of being very fast and generating outputs that do not require further processing (see section 2.1.4 and Annex B).

The way these components were implemented and packaged (e.g. high processing performance, user-friendly interface, automated segmentation parameter determination) resulted in a tool that offered an attractive middle ground between trivial methods and "cutting-edge" algorithms that are often computationally complex and do not generalise to real-world scenarios routinely encountered by cell culture researchers. The demand by the community for such tools was further demonstrated by the high number of downloads (estimated to be ~1500 as of December 2014) of the open-source release of these algorithms for MATLAB and FIJI. In an effort to reduce the barrier of entry for new users, the release of these algorithms was accompanied by a tool allowing the automated determination of near-optimal segmentation parameters using the Nelder-Mead optimisation method. No thorough characterisation of the optimisation problem was carried out for this work, which means that Nelder-Mead was unlikely to be the most appropriate optimisation method (see section 2.4.2). However, its speed and ease of implementation made it suitable for a proof-of-concept demonstration of automated segmentation parameter determination.

Whereas the contrast thresholding approach successfully enhanced the applicability of image processing for cell culture researchers (by combining well-understood principles in an novel fashion), the proposed trainable segmentation scheme constituted an original scientific contribution to the field of microscopy image processing and analysis in itself. The use of

machine learning techniques for pixel-wise segmentation of microscopy image was pioneered and made popular by Ilastik and the Weka trainable segmentation plugin for FIJI (Hall et al., 2009; Sommer et al., 2011b). These two software packages implement very similar approaches for pixel classification. First, image features (e.g. intensity, texture, edge map) are computed at multiple scales. For each pixel, the value obtained across all features and scales are then concatenated to form a pixel feature vector, based on which a Random Forest classifier predicts the pixel class label (e.g. cell or background). At the classification stage, pixels are effectively taken as independent entities; while their arrangement will dictate the value obtained for the image features considered, spatial information is mostly ignored. Random forest, in addition to being already successfully used by both Ilastik and Weka trainable segmentation, was chosen due to its low computational complexity, thus allowing the classification of millions of pixels based on hundreds of features in a timely fashion. Moreover, it was shown to have a good accuracy and to be less prone to overfitting than other classifiers (see section 4.1.2).

In the new scheme proposed, spatial context is encoded by computing local histograms at each pixel location for all combinations of features and scales considered. In order to make the computation of this large number of histograms tractable, a number of histogram construction methods were compared. The most efficient were found to be the integral histogram and computation by convolution approaches. In addition to generic image features such as intensity and local contrast, higher level image structure descriptors were also evaluated. While a large number of these descriptors have been proposed in the literature, this work employed Basic Image Features (BIFs) due to their ability to capture informative image patterns and structures in very compact representations that are particularly suitable for trainable segmentation (see section 4.5 for more information).

Using the proposed multi-scale local feature histogram scheme, moderate improvements were obtained for simple cell versus background segmentation tasks. However, drastic increases in segmentation performance (~1.5-fold) were observed when considering the more complex task of discriminating between different cell types. Based on these results, the proposed scheme was considered an improvement over the state-of-the-art for trainable segmentation of light microscopy images.

A MATLAB implementation of this novel trainable segmentation scheme is going to be made available under an open-source license. Moreover, the addition of this capability to both Ilastik and Weka trainable segmentation plugin is planned for the future.

Glossary

Binary image	Type of image where all elements take the values 0 (usually denoting background pixels) or 1 (usually denoting foreground pixels). Binary images (or masks) are usually obtained after segmentation, for example using a thresholding method.
Confluency	Fraction of the culture area occupied by cells. Confluency is used to qualitatively assess the state of a culture through visual inspection. *See section 3.1.1*
Contrast thresholding	Image segmentation method where the application of a contrast filter assigns high pixel values to regions of high intensity variations (i.e. cells) and low pixel values to homogeneous regions (i.e. image background) *See also intensity thresholding and section 2.1.2*
Decision trees	Type of machine learning algorithm where predictions are determined by a sequence of "if – then – else" statements termed nodes. At each node, a split function that applies a threshold on a subset of features dictates whether the left or right child node should be visited next. Predicted class is obtained by "pushing" the data down the tree. *See section 4.1.2*
Differentiation	Process during which cells acquire a specialised phenotype, usually accompanied by a loss in plasticity/potency. *See pluripotent stem cells*
Embryonic stem cells (ESCs)	*See pluripotent stem cells*
Halo artefact	Type of artefacts present in phase contrast microscopy images. Takes the form of rings made up of very bright pixels located around cellular objects.
Intensity thresholding	Image processing operations where all pixels of an image whose intensity value is above a given threshold will be set to 1, and 0 otherwise. The result is thus a binary image. *See also contrast thresholding and section 2.1.2*
Morphometric analysis	Characterisation of shape properties of objects in an image. Examples include the area, perimeter, convex hull, and form factor. *See section 3.1.2*
Packing corrected confluency (PCC)	Confluency corrected for changes in specific cell area that usually occur during the course of a culture for colony-forming cell lines. *See section 3.3*
Phase contrast microscopy	Type of light microscopy where variations in phase are transformed into variations in amplitude using an optics arrangement, so that phase specimen such as cells can be observed.
Pluripotent stem cells (PSCs)	Cells that can give rise to any adult cell type through differentiation. In this work, PSCs refer to either embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs).

Random (decisions) forest	Type of machine learning algorithm part of the ensemble family whose predictions are based on the majority vote across several decision trees (weak learners). It combines bagging (bootstrap aggregating) and random sampling of features at each node to produce low variance, low bias models. *See section 4.1.2*
Serial sacrificial experiment	Experimental design where multiple cultures are started simultaneously and then sequentially sacrificed for analysis following a pre-established schedule. Serial sacrificial experiments are usually used for the generation of time-course data when non-invasive determination of the variable of interest is not possible.
Shade-off effect	Type of artefacts present in phase contrast microscopy images that result in background and foreground objects having very similar intensity values (i.e. low contrast).
Single cell suspension	Suspension containing cells that are assumed to be well-separated (i.e. no clumps or aggregates), usually used for culture seeding. Adherent cells can be mechanically or enzymatically disaggregated and re-suspended to form a single cell suspension.
Specific cell area	The total area occupied by cells divided by the number of cells *See section 3.3*

Bibliography

Abaci HE, Truitt R, Luong E, Drazer G, Gerecht S. 2010. Adaptation to oxygen deprivation in cultures of human pluripotent stem cells, endothelial progenitor cells, and umbilical vein endothelial cells. *Am. J. Physiol. Cell Physiol.* **298**:C1527–37.

Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**:2274–82.

Ahmed T, Stocker R. 2008. Experimental verification of the behavioral foundation of bacterial transport parameters using microfluidics. *Biophys. J.* **95**:4481–93.

Albeanu DF, Soucy E, Sato TF, Meister M, Murthy VN. 2008. LED arrays as cost effective and efficient light sources for widefield microscopy. Ed. Rachel O. L. Wong. *PLoS One* **3**:e2146.

Albeck S, Nordlund P, Sussman JL, Nettleship JE, Assenberg R, Diprose JM, Rahman-Huq N, Owens RJ. 2010. Recent advances in the production of proteins in insect and mammalian cells for structural biology. *J. Struct. Biol.* **172**:55–65.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. Molecular biology of the cell 4th ed. Garland.

Alfred R, Taiani JT, Krawetz RJ, Yamashita A, Rancourt DE, Kallos MS. 2011. Large-scale production of murine embryonic stem cell-derived osteoblasts and chondrocytes on microcarriers in serum-free media. *Biomaterials* **32**:6006–6016.

Ali R, Gooding M, Szilágyi T, Vojnovic B, Christlieb M, Brady M. 2011. Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images. *Mach. Vis. Appl.* **23**:607–621.

Allen DD, Caviedes R, Cárdenas AM, Shimahara T, Segura-Aguilar J, Caviedes PA. 2005. Cell Lines as In Vitro Models for Drug Screening and Toxicity Studies. *Drug Dev. Ind. Pharm.* **31**:757–768.

Ambühl ME, Brepsant C, Meister J-J, Verkhovsky a B, Sbalzarini IF. 2012. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *J. Microsc.* **245**:161–70.

Antony J. 2003. Design of Experiments for Engineers and Scientists (Google eBook). Butterworth-Heinemann.

Bagnaninchi PO, Drummond N. 2011. Real-time label-free monitoring of adipose-derived stem cell differentiation with electric cell-substrate impedance sensing. *Proc. Natl. Acad. Sci. U. S. A.* **108**:6462–7.

Baharvand H, Azarnia M, Parivar K, Ashtiani SK. 2005. The effect of extracellular matrix on embryonic stem cell-derived cardiomyocytes. *J. Mol. Cell. Cardiol.* **38**:495–503.

Balagadde F, Song H, Ozaki J, Collins C, Barnet M, Arnold F, Quake S, You L. 2008. A synthetic Escherichia coli predator-prey ecosystem. *Mol Syst Biol* **4**:187.

Barbulovic-Nad I, Au SH, Wheeler AR. 2010. A microfluidic platform for complete mammalian cell culture. *Lab Chip* **10**:1536–1542.

Bawn A, Lee HW, Downey A, Xu J, Starkey JA, Yoon S. 2013. Metabolic-sensing characteristics of absorption-photometry for mammalian cell cultures in biopharmaceutical processes. *Pharm. Bioprocess.* **1**:255–266.

Bendall SC, Nolan GP, Roederer M, Chattopadhyay PK. 2012. A deep profiler's guide to cytometry. *Trends Immunol.* **33**:323–32.

Berger B, Peng J, Singh M. 2013. Computational solutions for omics data. *Nat. Rev. Genet.* **14**:333–46.

Bin Z, Sheng L, Gang Z, Hong J, Jun C, Bo Y, Hui S. 2006. Efficient cardiomyocyte differentiation of embryonic stem cells by bone morphogenetic protein-2 combined with visceral endoderm-like cells. *Cell Biol. Int.* **30**:769–776.

Bishop CM. 2006. Pattern Recognition and Machine Learning. New York, New York, USA: Springer New York.

Blanco-Gelaz MA, Suarez-Alvarez B, Ligero G, Sanchez L, Vidal-Castiñeira JR, Coto E, Moore H, Menendez P, Lopez-Larrea C. 2010. Endoplasmic reticulum stress signals in defined human embryonic stem cell lines and culture conditions. *Stem Cell Rev.* **6**:462–72.

Bradhurst CJ, Boles W, Xiao Y. 2008. Segmentation of bone marrow stromal cells in phase contrast microscopy images. In: . *23rd IVCNZ*. IEEE, pp. 1–6.

Breiman L. 1996. Bagging Predictors. *Mach. Learn.* **24**:123–140.

Breiman L. 2001a. Random Forests. Ed. Robert E Schapire. *Mach. Learn.* **45**. Springer Series in Statistics:5–32.

Breiman L. 2001b. Statistical Modeling: The Two Cultures. *Stat. Sci.* **16**:199–231.

Burnette WN. 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**:195–203.

Bustin SA. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **25**:169–93.

Cai AQ, Landman KA, Hughes BD. 2007. Multi-scale modeling of a wound-healing cell migration assay. *J. Theor. Biol.* **245**:576–94.

Canny J. 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**:679–98.

Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin D a, Chang JH, Lindquist R a, Moffat J, Golland P, Sabatini DM. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**:R100.

Carrel A. 1912. PURE CULTURES OF CELLS. *J. Exp. Med.* **16**:165–168.

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263**:802–5.

Chandler KJ, Barrier M, Jeffay S, Nichols HP, Kleinstreuer NC, Singh A V, Reif DM, Sipes NS, Judson RS, Dix DJ, Kavlock R, Hunter ES, Knudsen TB. 2011. Evaluation of 309 environmental chemicals using a mouse embryonic stem cell adherent cell differentiation and cytotoxicity assay. Ed. Austin John Cooney. *PLoS One* **6**:e18540.

Chazotte B. 2010. Labeling cytoskeletal F-actin with rhodamine phalloidin or fluorescein phalloidin for imaging. *Cold Spring Harb. Protoc.* **2010**:pdb.prot4947.

Chazotte B. 2011. Labeling membrane glycoproteins or glycolipids with fluorescent wheat germ agglutinin. *Cold Spring Harb. Protoc.* **2011**:pdb.prot5623.

Chen AK-L, Chen X, Choo ABH, Reuveny S, Oh SKW. 2011. Critical microcarrier properties affecting the expansion of undifferentiated human embryonic stem cells. *Stem Cell Res.* **7**:97–111.

Chen S-W, Yang JM, Yang J-H, Yang SJ, Wang J-S. 2012. A computational modeling and analysis in cell biological dynamics using electric cell-substrate impedance sensing (ECIS). *Biosens. Bioelectron.* **33**:196–203.

Cheng Y. 1995. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**:790–799.

Cheung Y-T, Lau WK-W, Yu M-S, Lai CS-W, Yeung S-C, So K-F, Chang RC-C. 2009. Effects of all-trans-retinoic acid on human SH-SY5Y neuroblastoma as in vitro model in neurotoxicity research. *Neurotoxicology* **30**:127–35.

Chin VI, Taupin P, Sanga S, Scheel J, Gage FH, Bhatia SN. 2004. Microfabricated platform for studying stem cell fates. *Biotechnol. Bioeng.* **88**:399–415.

Cho YM, Jang YS, Jang YM, Chung SM, Kim HS, Lee JH, Jeong SW, Kim IK, Kim JJ, Kim KS, Kwon OJ. 2009. Induction of unfolded protein response during neuronal induction of rat bone marrow stromal cells and mouse embryonic stem cells. *Exp. Mol. Med.* **41**:440–52.

Chu L, Robinson DK. 2001. Industrial choices for protein production by large-scale cell culture. *Curr. Opin. Biotechnol.* **12**:180–7.

Cireşan D, Meier U, Masci J, Schmidhuber J. 2012. Multi-column deep neural network for traffic sign classification. *Neural Netw.* **32**:333–8.

Cochran DM, Fukumura D, Ancukiewicz M, Carmeliet P, Jain RK. 2006. Evolution of oxygen and glucose concentration profiles in a tissue-mimetic culture system of embryonic stem cells. *Ann. Biomed. Eng.* **34**:1247–58.

Corish P, Tyler-Smith C. 1999. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng* **12**:1035–1040.

Criminisi A, Shotton J. 2013. Decision Forests for Computer Vision and Medical Image Analysis. Ed. Antonio Criminisi, J Shotton. Springer New York.

Crosier M, Griffin L. 2010. Using Basic Image Features for Texture Classification. *Int. J. Comput. Vis.* **88**:447–460.

Csaszar E, Kirouac DC, Yu M, Wang W, Qiao W, Cooke MP, Boitano AE, Ito C, Zandstra PW. 2012. Rapid expansion of human hematopoietic stem cells by automated control of inhibitory feedback signaling. *Cell Stem Cell* **10**:218–29.

Cui X, Lew M, Yang C. 2008. Quantitative differential interference contrast microscopy based on structured-aperture interference. *Appl. Phys. Lett.* **93**:091113.

Darling EM, Athanasiou KA. 2005. Rapid phenotypic changes in passaged articular chondrocyte subpopulations. *J. Orthop. Res.* **23**:425–32.

Darzynkiewicz Z, Holden E, Orfao A, Telford W, Wlodkowic D, Lin KK, Goodell MA. 2011. Detection of Hematopoietic Stem Cells by Flow Cytometry. In: . *Methods Cell Biol.*, Vol. 103, pp. 21–30.

Davila JC, Cezar GG, Thiede M, Strom S, Miki T, Trosko J. 2004. Use and application of stem cells in toxicology. *Toxicol. Sci.* **79**:214–223.

Debeir O, Adanja I, Warzee N, Van Ham P, Decaestecker C. 2008. Phase contrast image segmentation by weak watershed transform assembly. In: . *Biomed. Imaging From Nano to Macro, 2008. ISBI 2008. 5th IEEE Int. Symp.* IEEE, pp. 724–727.

Debeir O, Van Ham P, Kiss R, Decaestecker C. 2005. Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes. *IEEE Trans. Med. Imaging* **24**:697–711.

Deeks SG, McCune JM. 2010. Can HIV be cured with stem cell therapy? *Nat. Biotechnol.* **28**:807–810.

Dehlinger D, Suer L, Elsheikh M, Peña J, Naraghi-Arani P. 2013. Dye free automated cell counting and analysis. *Biotechnol. Bioeng.* **110**:838–47.

Denk W, Strickler JH, Webb WW. 1990. Two-photon laser scanning fluorescence microscopy. *Science* **248**:73–6.

Dodet B, Lechenet J, Hesselink W, Jungbäck C, Vannier P, Ottiger H-P, Huang L-C, Lin W, Yagami M, Tseng D, Miyashita-Lin E, Singh N, Lin A, Shih S-J. 2010. Validation of cell density and viability assays using Cedex automated cell counter. *Biologicals* **38**:393–400.

Dong JM, Leung T, Manser E, Lim L. 1998. cAMP-induced morphological changes are counteracted by the activated RhoA small GTPase and the Rho kinase ROKalpha. *J. Biol. Chem.* **273**:22554–62.

Doran PM. 2013. Bioprocess Engineering Principles Second Edi. Waltham: Academic Press.

Durand RE, Olive PL. 1982. Cytotoxicity, Mutagenicity and DNA damage by Hoechst 33342. *J. Histochem. Cytochem.* **30**:111–6.

Eibl R, Eibl D. 2011. Single-Use Technology in Biopharmaceutical Manufacture. John Wiley & Sons. Vol. 57.

Evan GI, d'Adda di Fagagna F. 2009. Cellular senescence: hot or what? *Curr. Opin. Genet. Dev.* **19**:25–31.

Fawcett T. 2006. An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**:861–874.

Figallo E, Cannizzaro C, Gerecht S, Burdick JA, Langer R, Elvassore N, Vunjak-Novakovic G. 2007. Micro-bioreactor array for controlling cellular microenvironments. *Lab Chip* **7**:710–719.

Fisher RA. 1971. The Design of Experiments. Macmillan Pub Co; 9 edition.

Fisher RA. 1992. The arrangement of field experiments. In: . *Break. Stat.*, Vol. 33, pp. 82–91.

Flight R, Landini G, Styles I, Shelton R, Milward M, Cooper P. 2014. Semi-automated cell counting in phase contrast images of epithelial monolayers. In: . *Med. Image Underst. Anal. 2014*, pp. 241–246.

Fowlkes C, Belongie S, Chung F, Malik J. 2004. Spectral grouping using the Nyström method. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**:214–25.

Freund Y. 2009. A more robust boosting algorithm. Machine Learning.

Freund Y, Schapire RE. 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **55**:119–139.

Fritzsche M, Mandenius C-F. 2010. Fluorescent cell-based sensing approaches for toxicity testing. *Anal. Bioanal. Chem.* **398**:181–91.

Fulwyler MJ. 1965. Electronic separation of biological cells by volume. *Science* **150**:910–1.

Galdeen SA, North AJ. 2011. Live cell fluorescence microscopy techniques. *Methods Mol. Biol.* **769**:205–22.

Gallagher R. 1999. Beyond Reductionism. *Science (80-.).* **284**:79–79.

Gao S, McGarry M, Ferrier T, Pallante B, Priddle H, Gasparrini B, Fletcher J, Harkness L, De Sousa P, McWhir J, Wilmut I. 2003. Effect of cell confluence on production of cloned mice using an inbred embryonic stem cell line. *Biol. Reprod.* **68**:595–603.

Gasser F, Mulsant P, Gillois M. 1985. Long-term multiplication of the Chinese hamster ovary (CHO) cell line in a serum-free medium. *In Vitro Cell. Dev. Biol.* **21**:588–592.

Genzel Y, Rödig J, Rapp E, Reichl U. 2014. Vaccine production: upstream processing with adherent or suspension cell lines. *Methods Mol. Biol.* **1104**:371–93.

Giaever I, Keese CR. 1993. A morphological biosensor for mammalian cells. *Nature* **366**:591–2.

Giaever I. 1991. Micromotion of Mammalian Cells Measured Electrically. *Proc. Natl. Acad. Sci.* **88**:7896–7900.

Giulitti S, Magrofuoco E, Prevedello L, Elvassore N. 2013. Optimal periodic perfusion strategy for robust long-term microfluidic cell culture. *Lab Chip* **13**:4430–41.

Goldman S. 2005. Stem and progenitor cell-based therapy of the human central nervous system. *Nat. Biotechnol.* **23**:862–871.

Gonzalez RC, Woods RE. 2007. Digital Image Processing (3rd Edition). *3nd Ed.* Prentice Hall.

Goto H, Yang B, Petersen D, Pepper KA, Alfaro PA, Kohn DB, Reynolds CP. 2003. Transduction of green fluorescent protein increased oxidative stress and enhanced sensitivity to cytotoxic drugs in neuroblastoma cell lines. *Mol. Cancer Ther.* **2**:911–917.

Gouveia RM, Castelletto V, Alcock SG, Hamley IW, Connon CJ. 2013. Bioactive films produced from self-assembling peptide amphiphiles as versatile substrates for tuning cell adhesion and tissue architecture in serum-free conditions. *J. Mater. Chem. B* **1**:6157.

Griffin L, Lillholm M, Crosier M. 2009. Scale Space and Variational Methods in Computer Vision. Ed. Xue-Cheng Tai, Knut Mørken, Marius Lysaker, Knut-Andreas Lie. *Scale Sp.* Berlin, Heidelberg: Springer Berlin Heidelberg. Vol. 5567. Lecture Notes in Computer Science.

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. 2009. The WEKA data mining software. *ACM SIGKDD Explor. Newsl.* **11**:10.

Hand a J, Sun T, Barber DC, Hose DR, MacNeil S. 2009. Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration. *J. Microsc.* **234**:62–79.

Hanko VP, Rohrer JS. 2004. Determination of amino acids in cell culture and fermentation broth media using anion-exchange chromatography with integrated pulsed amperometric detection. *Anal. Biochem.* **324**:29–38.

Hanson MA, Ge X, Kostov Y, Brorson KA, Moreira AR, Rao G. 2007. Comparisons of Optical pH and Dissolved Oxygen Sensors With Traditional Electrochemical Probes During Mammalian Cell Culture. *Biotechnology* **97**:833–841.

Haralick RM. 1979. Statistical and structural approaches to texture. *Proc. IEEE* **67**:786–804.

Hastie T, Tibshirani R, Friedman J. 2009. The Elements of Statistical Learning. *Math. ....* New York, NY: Springer New York. Springer Series in Statistics.

Hayashi Y, Furue MK, Okamoto T, Ohnuma K, Myoishi Y, Fukuhara Y, Abe T, Sato JD, Hata R-I, Asashima M. 2007. Integrins regulate mouse embryonic stem cell self-renewal. *Stem Cells* **25**:3005–3015.

Hayflick L. 1965. The limited in vitro lifetime of Human diploid cell strains. *Exp. Cell Res.* **37**:614–636.

Herbert M, Nayernia K, Keogh C. 2009. New Developments in Stem Cell Research. *Imaging Microsc.* **11**:29–31.

Hirst SJ, Twort CH, Lee TH. 2000. Differential effects of extracellular matrix proteins on human airway smooth muscle cell proliferation and phenotype. *Am. J. Respir. Cell Mol. Biol.* **23**:335–44.

Holzner C, Feser M, Vogt S, Hornberger B, Baines SB, Jacobsen C. 2010. Zernike phase contrast in scanning microscopy with X-rays. *Nat. Phys.* **6**:883–887.

Hu AY-C, Weng T-C, Tseng Y-F, Chen Y-S, Wu C-H, Hsiao S, Chou A-H, Chao H-J, Gu A, Wu S-C, Chong P, Lee M-S. 2008. Microcarrier-based MDCK cell culture system for the production of influenza H5N1 vaccines. *Vaccine* **26**:5736–40.

Jaalouk DE, Lammerding J. 2009. Mechanotransduction gone awry. *Nat. Rev. Mol. Cell Biol.* **10**:63–73.

Jamur MC, Oliver C. 2010. Permeabilization of cell membranes. *Methods Mol. Biol.* **588**:63–6.

Janakiraman V, Englert D, Jayaraman A, Baskaran H. 2009. Modeling growth and quorum sensing in biofilms grown in microfluidic chambers. *Ann. Biomed. Eng.* **37**:1206–16.

Jannin P. 2012. Validation in Medical Image Processing: Methodological Issues for Proper Quantification of Uncertainties. *Curr. Med. Imaging Rev.* **8**:322–330.

Jannin P, Krupinski E, Warfield SK, others. 2006. Validation in medical image processing. *IEEE Trans. Med. Imaging* **25**:1405–1409.

Jenson HB, Grant GM, Ench Y, Heard P, Thomas CA, Hilsenbeck SG, Moyer MP. 1998. Immunofluorescence microscopy and flow cytometry characterization of chemical induction of latent Epstein-Barr virus. *Clin. Diagn. Lab. Immunol.* **5**:91–7.

Jeong H-W, Nam J-O, Kim I-S. 2005. The COOH-terminal end of R-Ras alters the motility and morphology of breast epithelial cells through Rho/Rho-kinase. *Cancer Res.* **65**:507–15.

Jinno S, Lin J, Yageta M, Okayama H. 2001. Oncogenic cell cycle start control. *Mutat. Res. Mol. Mech. Mutagen.* **477**:23–29.

Johnson L V., Walsh ML, Chen LB. 1980. Localization of mitochondria in living cells with rhodamine 123. *Proc. Natl. Acad. Sci.* **77**:990–994.

Juneau P-M, Garnier A, Duchesne C. 2013. Selection and tuning of a fast and simple phase-contrast microscopy image segmentation algorithm for measuring myoblast growth kinetics in an automated manner. *Microsc. Microanal.* **19**:855–66.

Kamei K-I, Guo S, Yu ZTF, Takahashi H, Gschweng E, Suh C, Wang X, Tang J, McLaughlin J, Witte ON, Lee K-B, Tseng H-R. 2009. An integrated microfluidic culture device for quantitative analysis of human embryonic stem cells. *Lab Chip* **9**:555–563.

Kass M, Witkin A, Terzopoulos D. 1988. Snakes: Active contour models. *Int. J. Comput. Vis.* **1**:321–331.

Kato R, Iejima D, Agata H, Asahina I, Okada K, Ueda M, Honda H, Kagami H. 2010. A compact, automated cell culture system for clinical scale cell expansion from primary tissues. *Tissue Eng. Part C. Methods* **16**:947–56.

Kazmar T, Smid M, Fuchs M, Luber B, Mattes J. 2010. Learning cellular texture features in microscopic cancer cell images for automated cell-detection. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2010**:49–52.

Ker DFE, Weiss LE, Junkers SN, Chen M, Yin Z, Sandbothe MF, Huh S-I, Eom S, Bise R, Osuna-Highley E, Kanade T, Campbell PG. 2011. An engineered approach to stem cell culture: automating the decision process for real-time adaptive subculture of stem cells. *PLoS One* **6**:e27672.

Khetani SR, Bhatia SN. 2008. Microscale culture of human liver cells for drug development. *Nat. Biotechnol.* **26**:120–6.

Kim BJ, Wu M. 2012. Microfluidics for mammalian cell chemotaxis. *Ann. Biomed. Eng.* **40**:1316–27.

Kim JS, Hur D, Hwang JK, Chung C, Chang JK. 2010. Ongoing development of image cytometers. *Bioanalysis* **2**:1755–65.

Kim L, Toh Y-C, Voldman J, Yu H. 2007. A practical guide to microfluidic perfusion culture of adherent mammalian cells. *Lab Chip* **7**:681–94.

Kim L, Vahey MD, Lee H-Y, Voldman J. 2006. Microfluidic arrays for logarithmically perfused embryonic stem cell culture. *Lab Chip* **6**:394–406.

King KR, Wang S, Irimia D, Jayaraman A, Toner M, Yarmush ML. 2007. A high-throughput microfluidic real-time gene expression living cell array. *Lab Chip* **7**:77–85.

Kirsch RA. 1971. Computer determination of the constituent structure of biological images. *Comput. Biomed. Res.* **4**:315–28.

Kitano H. 2002. Systems biology: a brief overview. *Science* **295**:1662–4.

Kohl P, Crampin EJ, Quinn TA, Noble D. 2010. Systems biology: an approach. *Clin. Pharmacol. Ther.* **88**:25–33.

Korin N, Bransky A, Dinnar U, Levenberg S. 2009. Periodic "flow-stop" perfusion microchannel bioreactors for mammalian and human embryonic stem cell long-term culture. *Biomed. Microdevices* **11**:87–94.

Kretlow JD, Jin Y-Q, Liu W, Zhang WJ, Hong T-H, Zhou G, Baggett LS, Mikos AG, Cao Y. 2008. Donor age and cell passage affects differentiation potential of murine bone marrow-derived stem cells. *BMC Cell Biol.* **9**:60.

Kühn J, Shaffer E, Mena J, Breton B, Parent J, Rappaz B, Chambon M, Emery Y, Magistretti P, Depeursinge C, Marquet P, Turcatti G. 2013. Label-free cytotoxicity screening assay by digital holographic microscopy. *Assay Drug Dev. Technol.* **11**:101–7.

Kuska J-P, Scheibe P, Braumann U-D. 2008. Fast fluid extensions for image registration algorithms. In: . *2008 15th IEEE Int. Conf. Image Process.* IEEE, pp. 2408–2411.

Lakowicz JR. 2010. Principles of Fluorescence Spectroscopy. Springer.

Lampugnani MG, Corada M, Andriopoulou P, Esser S, Risau W, Dejana E. 1997. Cell confluence regulates tyrosine phosphorylation of adherens junction components in endothelial cells. *J. Cell Sci.* **110 (Pt 1**:2065–77.

Langdon SP. 2004. Basic principles of cancer cell culture. *Methods Mol. Med.* **88**:3–15.

Lantz RC. 2001. Rapid Reduction of Intracellular Glutathione in Human Bronchial Epithelial Cells Exposed to Occupational Levels of Toluene Diisocyanate. *Toxicol. Sci.* **60**:348–355.

Lawrence S. 1852. The inverted microscope-a new form of microscope. *Am. J. Sci. arts* **14**:233–241.

LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. 1989. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* **1**:541–551.

Lecun Y, Bottou L, Bengio Y, Haffner P. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**:2278–2324.

Leung-Hagesteijn C, Milankov K, Michalak M, Wilkins J, Dedhar S. 1994. Cell attachment to extracellular matrix substrates is inhibited upon downregulation of expression of calreticulin, an intracellular integrin alpha-subunit-binding protein. *J. Cell Sci.* **107**:589–600.

Lew C, Gomez JA, Rhyner MN. 2012. Instrument-to-instrument Variability in the Vi-CELL Automated Viability Analyzer. https://www.beckmancoulter.com/wsrportal/bibliography?docname=IB-17279.pdf.

Li CLC, Xu CXC, Gui CGC, Fox MD. 2005. Level set evolution without re-initialization: a new variational formulation. *2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **1**.

Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, Huang CC, Kain SR. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* **273**:34970–5.

Liu C-L, Nakashima K, Sako H, Fujisawa H. 2003. Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern Recognit.* **36**:2271–2285.

Liu HS, Jan MS, Chou CK, Chen PH, Ke NJ. 1999. Is green fluorescent protein toxic to the living cells? *Biochem. Biophys. Res. Commun.* **260**:712–717.

Lulevich V, Shih Y-P, Lo SH, Liu G-Y. 2009. Cell tracing dyes significantly change single cell mechanics. *J. Phys. Chem. B* **113**:6511–9.

Lundholt BK, Scudder KM, Pagliaro L. 2003. A simple technique for reducing edge effect in cell-based assays. *J. Biomol. Screen.* **8**:566–70.

Lyons a B. 1999. Divided we stand: tracking cell proliferation with carboxyfluorescein diacetate succinimidyl ester. *Immunol. Cell Biol.* **77**:509–15.

Mak GW-Y, Wong C-H, Tsui SK-W. 2007. Green fluorescent protein induces the secretion of inflammatory cytokine interleukin-6 in muscle cells. *Anal. Biochem.* **362**:296–8.

Marquet P, Depeursinge C, Magistretti PJ. 2013. Exploring neural cell dynamics with digital holographic microscopy. *Annu. Rev. Biomed. Eng.* **15**:407–31.

Mehling M, Tay S. 2014. Microfluidic cell culture. *Curr. Opin. Biotechnol.*

Millet LJ, Stewart ME, Nuzzo RG, Gillette MU. 2010. Guiding neuron development with planar surface gradients of substrate cues deposited using microfluidic devices. *Lab Chip* **10**:1525–35.

Mimeault M, Hauke R, Batra SK. 2007. Stem cells: a revolution in therapeutics-recent advances in stem cell biology and their therapeutic applications in regenerative medicine and cancer therapies. *Clin. Pharmacol. Ther.* **82**:252–264.

Mölder A, Sebesta M, Gustafsson M, Gisselson L, Wingren AG, Alm K. 2008. Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography. *J. Microsc.* **232**:240–7.

Murphy DB, Spring KR, Parry-Hill M, Davidson MW. 2013. Microscopyu: Shade-Off and Halo Phase Contrast Artifacts. http://www.microscopyu.com/tutorials/java/phasecontrast/shadeoff/.

Naciri M, Kuystermans D, Al-Rubeai M. 2008. Monitoring pH and dissolved oxygen in mammalian cell culture using optical sensors. *Cytotechnology* **57**:245–50.

Nam S-W, Van Noort D, Yang Y, Park S. 2007. A biological sensor platform using a pneumatic-valve controlled microfluidic device containing Tetrahymena pyriformis. *Lab Chip* **7**:638–40.

Narkilahti S, Rajala K, Pihlajamaki H, Suuronen R, Hovatta O, Skottman H. 2007. Monitoring and analysis of dynamic growth of human embryonic stem cells: comparison of automated instrumentation and conventional culturing methods. *Biomed Eng Online* **6**:11.

Nelder JA, Mead R. 1965. A simplex method for function minimization. *Comput. J.* **7**:308–313.

Nethercott HE, Brick DJ, Schwartz PH. 2011. Derivation of induced pluripotent stem cells by lentiviral transduction. *Methods Mol. Biol.* **767**:67–85.

Neuži P, Giselbrecht S, Länge K, Huang TJ, Manz A. 2012. Revisiting lab-on-a-chip technology for drug discovery. *Nat. Rev. Drug Discov.*

Newell AJ, Griffin LD. 2011. Natural Image Character Recognition Using Oriented Basic Image Features. *2011 Int. Conf. Digit. Image Comput. Tech. Appl.*:191–196.

Ng AYJ, Rajapakse JC, Welsch RE, Matsudaira PT, Horodincu V, Evans JG. 2010. A cell profiling framework for modeling drug responses from HCS imaging. *J. Biomol. Screen.* **15**:858–68.

Niswender KD, Blackman SM, Rohde L, Magnuson M a, Piston DW. 1995. Quantitative imaging of green fluorescent protein in cultured cells: comparison of microscopic techniques, use in fusion proteins and detection limits. *J. Microsc.* **180**:109–16.

Notingher I, Bisson I, Bishop AE, Randle WL, Polak JMP, Hench LL. 2004. In situ spectral monitoring of mRNA translation in embryonic stem cells during differentiation in vitro. *Anal. Chem.* **76**:3185–93.

Notingher I, Hench LL. 2006. Raman microspectroscopy: a noninvasive tool for studies of individual living cells in vitro. *Expert Rev. Med. Devices* **3**:215–34.

Odell ID, Cook D. 2013. Immunofluorescence techniques. *J. Invest. Dermatol.* **133**:e4.

Odorico J, Kaufman D, Thomson J. 2001. Multilineage differentiation from human embryonic stem cell lines. *Stem Cells* **19**:193–204.

Ojala T, Pietikäinen M, Harwood D. 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **29**:51–59.

Otaki T. 2000. Artifact Halo Reduction in Phase Contrast Microscopy Using Apodization. *Opt. Rev.* **7**:119–122.

Paddock SW. 2000. Principles and practices of laser scanning confocal microscopy. *Mol. Biotechnol.* **16**:127–49.

Pan J, Kanade T, Chen M. 2009. Learning to Detect Different Types of Cells under Phase Contrast Microscopy. In: . *Microsc. Image Anal. with Appl. Biol.*

Panchuk-Voloshina N, Haugland RP, Bishop-Stewart J, Bhalgat MK, Millard PJ, Mao F, Leung W-Y. 1999. Alexa Dyes, a Series of New Fluorescent Dyes that Yield Exceptionally Bright, Photostable Conjugates. *J. Histochem. Cytochem.* **47**:1179–1188.

Park Y Bin, Kim YY, Oh SK, Chung SG, Ku SY, Kim SH, Choi YM, Moon SY. 2008. Alterations of proliferative and differentiation potentials of human embryonic stem cells during long-term culture. *Exp. Mol. Med.* **40**:98–108.

Pascut FC, Goh HT, Welch N, Buttery LD, Denning C, Notingher I. 2011. Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells. *Biophys. J.* **100**:251–9.

Pasquier J, Galas L, Boulangé-Lecomte C, Rioult D, Bultelle F, Magal P, Webb G, Le Foll F. 2012. Different modalities of intercellular membrane exchanges mediate cell-to-cell p-glycoprotein transfers in MCF-7 breast cancer cells. *J. Biol. Chem.* **287**:7374–87.

Perry TL, Stedman D, Hansen S. 1968. A versatile lithium buffer elution system for single column automatic amino acid chromatography. *J. Chromatogr. A* **38**:460–466.

Petiot E, Fournier F, Gény C, Pinton H, Marc A. 2010. Rapid screening of serum-free media for the growth of adherent Vero cells by using a small-scale and non-invasive tool. *Appl. Biochem. Biotechnol.* **160**:1600–15.

Pierscionek BK, Li Y, Schachar RA, Chen W. 2012. The effect of high concentration and exposure duration of nanoceria on human lens epithelial cells. *Nanomedicine Nanotechnology, Biol. Med.* **8**:383–390.

Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, Bell SM, Combaret V, Puisieux A, Mighell AJ, Robinson PA, Inglehearn CF, Isaacs JD, Markham AF. 2003. Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnol.* **3**:18.

Porikli F. 2005. Integral histogram: a fast way to extract histograms in Cartesian spaces. In: . *2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* IEEE, Vol. 1, pp. 829–836 vol. 1.

Portugal J, Waring MJ. 1988. Assignment of DNA binding sites for 4',6-diamidine-2-phenylindole and bisbenzimide (Hoechst 33258). A comparative footprinting study. *Biochim. Biophys. Acta* **949**:158–68.

Pouton CW, Haynes JM. 2007. Embryonic stem cells as a source of models for drug discovery. *Nat. Rev. Drug Discov.* **6**:605–16.

Pozarowski P, Darzynkiewicz Z. 2004. Analysis of cell cycle by flow cytometry. *Methods Mol. Biol.* **281**:301–11.

Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ. 1992. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**:229–33.

Pretorius AJ, Bray M-AP, Carpenter AE, Ruddle RA. 2011. Visualization of Parameter Space for Image Analysis. *IEEE Trans. Vis. Comput. Graph.* **17**:2402–2411.

Puppels GJ, De Mul FF, Otto C, Greve J, Robert-Nicoud M, Arndt-Jovin DJ, Jovin TM. 1990. Studying single living cells and chromosomes by confocal Raman microspectroscopy. *Nature* **347**:301–303.

Pyrgiotakis G, Kundakcioglu OE, Finton K, Pardalos PM, Powers K, Moudgil BM. 2009. Cell death discrimination with Raman spectroscopy and support vector machines. *Ann. Biomed. Eng.* **37**:1464–73.

Reichen M, Macown RJ, Jaccard N, Super A, Ruban L, Griffin LD, Veraitch FS, Szita N. 2012. Microfabricated Modular Scale-Down Device for Regenerative Medicine Process Development. Ed. Costanza Emanueli. *PLoS One* **7**:e52246.

Reichen M, Veraitch FS, Szita N. 2013. Development of a multiplexed microfluidic platform for the automated cultivation of embryonic stem cells. *J. Lab. Autom.* **18**:519–29.

Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. 2000. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat. Biotechnol.* **18**:399–404.

Reubinoff B, Itsykson P, Turetsky T. 2001. Neural progenitors from human embryonic stem cells. *Nat. ...* **19**.

Richards M, Fong C-Y, Chan W-K, Wong P-C, Bongso A. 2002. Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells. *Nat. Biotechnol.* **20**:933–936.

Robinson GS. 1977. Edge detection by compass gradient masks. *Comput. Graph. Image Process.* **6**:492–501.

Roerdink JBTM, Meijster A. 2000. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. *Fundam. Informaticae* **41**:187–228.

Rudolph G, Lindner P, Gierse A, Bluma A, Martinez G, Hitzmann B, Scheper T. 2008. Online monitoring of microcarrier based fibroblast cultivations with in situ microscopy. *Biotechnol. Bioeng.* **99**:136–145.

Ruoslahti E, Reed JC. 1994. Anchorage dependence, integrins, and apoptosis. *Cell* **77**:477–478.

Ruutu M, Johansson B, Grenman R, Syrjänen K, Syrjänen S. 2004. Effect of confluence state and passaging on global cancer gene expression pattern in oral carcinoma cell lines. *Anticancer Res.* **24**:2627–31.

Ruzin SE. 1999. Plant Microtechnique and Microscopy. Oxford University Press.

Sahoo PK. 1990. Circularly symmetric separable functions are Gaussian. *Appl. Math. Lett.* **3**:111–113.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487–91.

Sakamoto N, Saito N, Han X, Ohashi T, Sato M. 2010. Effect of spatial gradient in fluid shear stress on morphological changes in endothelial cells in response to flow. *Biochem. Biophys. Res. Commun.* **395**:264–9.

Sartipy P, Björquist P, Strehl R, Hyllner J. 2007. The application of human embryonic stem cell technologies to drug discovery. *Drug Discov. Today* **12**:688–699.

Sato T, Murthy VN. 2012. Light-emitting diodes for biological microscopy. *Cold Spring Harb. Protoc.* **2012**:pdb.top072306–.

Scherf N, Herberg M, Thierbach K, Zerjatke T, Kalkan T, Humphreys P, Smith A, Glauche I, Roeder I. 2012. Imaging, quantification and visualization of spatio-temporal patterning in mESC colonies under different culture conditions. *Bioinformatics* **28**:i556–i561.

Schierbeek A, van Leeuwenhoek A, Rooseboom M. 1959. Measuring the Invisible World: The Life and Works of Antoni Van Leeuwenhoek FRS. Abelard-Schuman.

Schneider C a, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**:671–5.

Schoepflin T. 2001. Active contour model with gradient directional information: directional snake. *IEEE Trans. Circuits Syst. Video Technol.* **11**:252–256.

Segers VFM, Lee RT. 2008. Stem-cell therapy for cardiac disease. *Nature* **451**:937–942.

Selinummi J, Ruusuvuori P, Podolsky I, Ozinsky A, Gold E, Yli-Harja O, Aderem A, Shmulevich I. 2009. Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images. *PLoS One* **4**:e7497.

Seo S, Su T-W, Tseng DK, Erlinger A, Ozcan A. 2009. Lensfree holographic imaging for on-chip cytometry and diagnostics. *Lab Chip* **9**:777–87.

Seroussi I, Veikherman D, Ofer N, Yehudai-Resheff S, Keren K. 2012. Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes. *J. Microsc.* **247**:137–46.

Seyedhosseini M, Kumar R, Jurrus E, Giuly R, Ellisman M, Pfister H, Tasdizen T. 2011. Detection of neuron membranes in electron microscopy images using multi-scale context and radon-like features. *Med. Image Comput. Comput. Assist. Interv.* **14**:670–7.

Shakes DC, Miller DM, Nonet ML. 2012. Immunofluorescence microscopy. *Methods Cell Biol.* **107**:35–66.

Shaner NC, Steinbach PA, Tsien RY. 2005. A guide to choosing fluorescent proteins. *Nat. Methods* **2**:905–9.

SHIMOMURA O, JOHNSON FH, SAIGA Y. 1962. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* **59**:223–39.

Sison-Young RLC, Kia R, Heslop J, Kelly L, Rowe C, Cross MJ, Kitteringham NR, Hanley N, Park BK, Goldring CEP. 2012. Human pluripotent stem cells for modeling toxicity. *Adv. Pharmacol.* **63**:207–56.

Smith JR, Ning Y, Pereira-Smith OM. 1992. Why are transformed cells immortal? Is the process reversible? *Am. J. Clin. Nutr.* **55**:1215S–1221S.

Soltys Z, Orzylowska-Sliwinska O, Zaremba M, Orlowski D, Piechota M, Fiedorowicz a, Janeczko K, Oderfeld-Nowak B. 2005. Quantitative morphological study of microglial cells in the ischemic rat brain using principal component analysis. *J. Neurosci. Methods* **146**:50–60.

Sommer C, Straehle C, Kothe U, Hamprecht FA. 2011a. Ilastik: Interactive learning and segmentation toolkit. *2011 IEEE Int. Symp. Biomed. Imaging From Nano to Macro*:230–233.

Sommer C, Straehle C, Kothe U, Hamprecht FA. 2011b. Ilastik: Interactive learning and segmentation toolkit. In: . *2011 IEEE Int. Symp. Biomed. Imaging From Nano to Macro*. IEEE, pp. 230–233.

Stampfer MR, Bodnar A, Garbe J, Wong M, Pan A, Villeponteau B, Yaswen P. 1997. Gradual phenotypic conversion associated with immortalization of cultured human mammary epithelial cells. *Mol. Biol. Cell* **8**:2391–405.

Straub M, Lodemann P, Holroyd P, Jahn R, Hell SW. 2000. Live cell imaging by multifocal multiphoton microscopy. *Eur. J. Cell Biol.* **79**:726–34.

Strober W. 2001. Monitoring cell growth. *Curr. Protoc. Immunol.* **Appendix 3**:Appendix 3A.

Stroka KM, Vaitkus J a, Aranda-Espinoza H. 2012. Endothelial cells undergo morphological, biomechanical, and dynamic changes in response to tumor necrosis factor-α. *Eur. Biophys. J.* **41**:939–47.

Su H, Yin Z, Huh S, Kanade T. 2013a. Cell segmentation via spectral analysis on phase retardation features. In: . *2013 IEEE 10th Int. Symp. Biomed. Imaging*. IEEE, pp. 1477–1483.

Su H, Yin Z, Huh S, Kanade T. 2013b. Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features. *Med. Image Anal.* **17**:746–765.

Sugimoto N, Shido O, Matsuzaki K, Ohno-Shosaku T, Hitomi Y, Tanaka M, Sawaki T, Fujita Y, Kawanami T, Masaki Y, Okazaki T, Nakamura H, Koizumi S, Yachie A, Umehara H. 2012. Cellular heat acclimation regulates cell growth, cell morphology, mitogen-activated protein kinase activation, and expression of aquaporins in mouse fibroblast cells. *Cell. Physiol. Biochem.* **30**:450–7.

Sutherland FWH, Perry TE, Yu Y, Sherwood MC, Rabkin E, Masuda Y, Garcia GA, McLellan DL, Engelmayr GC, Sacks MS, Schoen FJ, Mayer JE. 2005. From stem cells to viable autologous semilunar heart valve. *Circulation* **111**:2783–91.

Tabas I, Ron D. 2011. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nat. Cell Biol.* **13**:184–90.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**:663–676.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**:861–872.

Theriault DH, Walker ML, Wong JY, Betke M. 2011. Cell morphology classification and clutter mitigation in phase-contrast microscopy images using machine learning. *Mach. Vis. Appl.* **23**:659–673.

Thomas RJ, Hourd PC, Williams DJ. 2008. Application of process quality engineering techniques to improve the understanding of the in vitro processing of stem cells for therapeutic use. *J. Biotechnol.* **136**:148–55.

Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Becker RA, Hearn JP. 1995. Isolation of a primate embryonic stem cell line. *Proc. Natl. Acad. Sci. U. S. A.* **92**:7844–7848.

Thomson J a., Itskovitz-Eldor J, Shapiro S, Waknitz M, Swiergiel J, Marshall V, Jones J. 1998. Embryonic stem cell lines derived from human blastocysts. *Sci. (New York, NY)* **282**:1145–1147.

Toh Y-C, Voldman J. 2011. Fluid shear stress primes mouse embryonic stem cells for differentiation in a self-renewing environment via heparan sulfate proteoglycans transduction. *FASEB J.* **25**:1208–17.

Topman G, Sharabani-Yosef O, Gefen A. 2011. A Method for Quick, Low-Cost Automated Confluency Measurements. *Microsc. Microanal.* **17**:915–922.

Usaj M, Torkar D, Kanduser M, Miklavcic D. 2011. Cell counting tool parameters optimization approach for electroporation efficiency determination of attached cells in phase contrast images. *J. Microsc.* **241**:303–14.

Vedaldi A, Fulkerson B. 2010. Vlfeat: an open and portable library of computer vision algorithms. In: . *Proc. Int. Conf. Multimed. - MM '10*. New York, New York, USA: ACM Press, p. 1469.

Veraitch FS, Scott R, Wong J-W, Lye GJ, Mason C. 2008. The impact of manual processing on the expansion and directed differentiation of embryonic stem cells. *Biotechnol. Bioeng.* **99**:1216–1229.

Villa-Diaz LG, Torisawa Y, Uchida T, Ding J, Nogueira-de-Souza NC, O'Shea KS, Takayama S, Smith GD. 2009. Microfluidic culture of single human embryonic stem cell colonies. *Lab Chip* **9**:1749–1755.

Viola P, Jones M. 2002. Robust real-time object detection. *Int. J. Comput. Vis.*

Vladisavljević GT, Khalid N, Neves MA, Kuroiwa T, Nakajima M, Uemura K, Ichikawa S, Kobayashi I. 2013. Industrial lab-on-a-chip: Design, applications and scale-up for drug discovery and delivery. *Adv. Drug Deliv. Rev.*

Waddington CH. 1957. The strategy of the genes: a discussion of some aspects of theoretical biology. Allen & Unwin.

Wang X-Y, Wang T, Bu J. 2011. Color image segmentation using pixel wise support vector machine classification. *Pattern Recognit.* **44**:777–787.

Wang Z, Kim M-C, Marquez M, Thorsen T. 2007. High-density microfluidic arrays for cell cytotoxicity analysis. *Lab Chip* **7**:740–5.

Want AJ, Nienow AW, Hewitt CJ, Coopman K. 2012. Large-scale expansion and exploitation of pluripotent stem cells for regenerative medicine purposes: beyond the T flask. *Regen. Med.* **7**:71–84.

Wegener J, Keese CR, Giaever I. 2000. Electric Cell–Substrate Impedance Sensing (ECIS) as a Noninvasive Means to Monitor the Kinetics of Cell Spreading to Artificial Surfaces. *Exp. Cell Res.* **259**:158–166.

Wei Y, Tao L. 2010. Efficient histogram-based sliding window. In: . *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3003–3010.

Weston SA, Parish CR. 1990. New fluorescent dyes for lymphocyte migration studies. *J. Immunol. Methods* **133**:87–97.

Wilson CA, Theriot JA. 2006. A correlation-based approach to calculate rotation and translation of moving cells. *IEEE Trans. Image Process.* **15**:1939–1951.

Wilson T. 2011. Resolution and optical sectioning in the confocal microscope. *J. Microsc.* **244**:113–21.

Wobus AM, Boheler KR. 2005. Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiol. Rev.* **85**:635–678.

Wu M-H, Chang Y-H, Liu Y-T, Chen Y-M, Wang S-S, Wang H-Y, Lai C-S, Pan T-M. 2011. Development of high throughput microfluidic cell culture chip for perfusion 3-dimensional cell culture-based chemosensitivity assay. *Sensors Actuators B Chem.* **155**:397–407.

Xiao C, Luong JHT. 2003. On-line monitoring of cell growth and cytotoxicity using electric cell-substrate impedance sensing (ECIS). *Biotechnol. Prog.* **19**:1000–5.

Xu C, Prince JL. 1998. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* **7**:359–69.

Xu W, Jericho MH, Meinertzhagen IA, Kreuzer HJ. 2001. Digital in-line holography for biological applications. *Proc. Natl. Acad. Sci. U. S. A.* **98**:11301–5.

Yao L, Bestwick CS, Bestwick LA, Maffulli N, Aspden RM. 2006. Phenotypic drift in human tenocyte culture. *Tissue Eng.* **12**:1843–9.

Yin Z, Bise R, Chen M, Kanade T. 2010. Cell segmentation in microscopy imagery using a bag of local Bayesian classifiers. In: . *2010 IEEE Int. Symp. Biomed. Imaging From Nano to Macro*. IEEE, Vol. 1, pp. 125–128.

Yin Z, Kanade T, Chen M. 2012. Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation. *Med. Image Anal.* **16**:1047–62.

Ying QL, Stavridis M, Griffiths D, Li M, Smith A. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.* **21**:183–186.

Yoshida H. 2007. ER stress and diseases. *FEBS J.* **274**:630–58.

Yoshimitsu R, Hattori K, Sugiura S, Kondo Y, Yamada R, Tachikawa S, Satoh T, Kurisaki A, Ohnuma K, Asashima M, Kanamori T. 2013. Microfluidic perfusion culture of human induced pluripotent stem cells under fully defined culture conditions. *Biotechnol. Bioeng.*

Young EWK, Beebe DJ. 2010. Fundamentals of microfluidic cell culture in controlled microenvironments. *Chem. Soc. Rev.* **39**:1036–48.

Zemel A, Rehfeldt F, Brown AEX, Discher DE, Safran SA. 2010. Optimal matrix rigidity for stress fiber polarization in stem cells. *Nat. Phys.* **6**:468–473.

Zernike F. 1942. Phase contrast, a new method for the microscopic observation of transparent objects part II. *Physica* **9**:974–986.

Zhang C, van Noort D. 2011. Cells in microfluidics. *Top. Curr. Chem.* **304**:295–321.

Zink D, Sadoni N, Stelzer E. 2003. Visualizing chromatin and chromosomes in living cells. *Methods* **29**:42–50.

Ziółkowska K, Kwapiszewski R, Brzózka Z. 2011. Microfluidic devices as tools for mimicking the in vivo environment. *New J. Chem.* **35**:979.

Annexes

A. Digital microscopy images

Digital images are discretised representation of continuous objects, both in space and in amplitude (Gonzalez and Woods, 2007). Spatial discretization is done through sampling of the spatial domain at discrete intervals. A digital image *I* can be represented as 2-dimensional or 3-dimensional arrays, for grayscale and colour images respectively. Elements of *I* are called pixels and their value corresponds to the intensity of the sampled regions of the image. Pixels are addressed by their spatial coordinates *x* and *y*, which are integers. By convention, a right-handed Cartesian coordinate system is used (i.e. the origin is at the top-left corner of an image).

Discretization of the amplitude (or intensity) is done through quantization (Gonzalez and Woods, 2007). The number of discrete intensity levels *L* is usually expressed as a power of 2.

$$L = 2^k \tag{A.1}$$

where *k* is the number of bits of the image. For example, pixels of an 8-bit image can take 2^8 distinct intensity values. Different data types can be used to store and represent the image, regardless of the number of discrete levels of *I*. For this work, all images were converted to a double-precision floating-point representation (values ranging from 0, black, to 1, white). The most basic representation is a binary image, where each pixel is either 0 or 1, usually corresponding to background and foreground (or object) pixels respectively.

B. Methods for the determination of the intensity gradient of a digital image

First, the intensity gradient of the image had to be determined. Given an hypothetical continuous image *f(x,y)*, the gradient at the point *(x,y)* is defined as:

$$\nabla f\left(x,y\right) = \left[\, G_x G_y \,\right] = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} \tag{A.2}$$

where G_x and G_y are the gradient, $\frac{\partial f}{\partial x}$ the difference in intensity in the x direction and $\frac{\partial f}{\partial y}$ the difference in intensity in the y direction. The magnitude of the gradient vector at the point *(x,y)* is thus given by:

$$\overline{\nabla f} = \sqrt{G_x^2 + G_y^2} \tag{A.3}$$

The direction of the gradient in relation to the x axis can be expressed as:

$$\alpha = \arctan\left(\frac{G_x}{G_y}\right) = \arctan\left(\frac{\partial f}{\partial y}\cdot\frac{\partial x}{\partial f}\right) \quad \text{A.4}$$

The gradient points in the direction of the most rapid change in pixel intensity. It is usually perpendicular to edges.

B.1.1 Using the finite difference method

As a digital image is by nature discrete, the computation of the intensity gradient is done by approximation using the central (or symmetric) finite difference. The derivative at the pixel *(x,y)* in the x direction is thus given by:

$$\frac{\partial I(x,y)}{\partial x} \approx \frac{I(x+n,y)-I(x-n,y)}{2n} \quad \text{A.5}$$

and in the y direction by:

$$\frac{\partial I(x,y)}{\partial y} \approx \frac{I(x,y+n)-I(x,y-n)}{2n} \quad \text{A.6}$$

For pixels on the edges of an image, the forward and backward differences are used accordingly.

B.1.2 Using convolution

Convolution of an image *I* with a kernel *H* is defined as:

$$c(x,y) = I(x,y)\otimes H(x,y) = \sum_{i=-n}^{n}\sum_{j=-n}^{n} H(i,j) * I(x-i,y-j) \quad \text{A.7}$$

The central difference equation for the horizontal direction can be rewritten as follows:

$$\frac{\partial I(x,y)}{\partial x} \approx -\frac{1}{2n} * I(x-n,y) + 0 * I(x,y) + \frac{1}{2n} * I(x+n,y) \quad \text{A.8}$$

Which is equivalent to the convolution of the image *I* by a kernel $H = \frac{1}{2n} * \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$:

$$\frac{\partial I(x,y)}{\partial x} \approx I(x,y)\otimes \frac{1}{2n} * \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad \text{A.9}$$

Similarly for the vertical derivative:

$$\frac{\partial I(x,y)}{\partial y} \approx I(x,y) \otimes \frac{1}{2n} * \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \tag{A.10}$$

This method however, is sensitive to image noise. It can be made more robust by combining it with an averaging (or smoothing) kernel:

$$H_x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} * \frac{1}{2n} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} = \frac{1}{2n} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \tag{A.11}$$

And for the vertical direction:

$$H_y = \frac{1}{2n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \frac{1}{2n} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \tag{A.12}$$

Many useful kernels for gradient determination and edge detection have been described in the literature, including Sobel and Prewitt kernels (Gonzalez and Woods, 2007). A variant of this technique termed Canny edge detection employed a Gaussian blur prior to edge detection in order to minimise the effect of noise (Canny, 1986).

Based on this principle, kernels that are sensitive to each of the 8 cardinal directions (often termed compass kernels) can be generated from templates that are sequentially rotated by 45°. Kirsch kernels (or compass operators) are one such example (Kirsch, 1971):

Kirsch kernels

$g^1(0°)$			$g^2(45°)$			$g^3(90°)$				$g^8(315°)$		
−3	−3	5	−3	5	5	5	5	5		−3	−3	−3
−3	0	5	−3	0	5	−3	0	−3	...	−3	0	5
−3	−3	5	−3	−3	−3	−3	−3	−3		−3	5	5

A.13

Other collection of rotated kernels that can be used for compass edge detection include Sobel and Robinson operators (Robinson, 1977).

The image is convolved with each of the kernels in the collection. The magnitude of the gradient is given by the maximum response:

$$c(x,y) = \max\left\{ \left| I(x,y) \otimes g^{(k)}(x,y) \right|, k = 1,2,...,8 \right\} \tag{A.14}$$

The direction of the gradient at that point is then given by the kernel that produced the maximum response. This method such has the advantage to directly output a gradient direction without the need to resort to quantisation from the gradient direction vectors.

B.1.3 Using convolution with a Gaussian kernel

As discussed above, differentiation is equivalent to convolution. Given that convolution is commutative and associative, for any image *I(x,y)* and kernel *H(x,y)* we can write:

$$\frac{\partial}{\partial x}\left[I(x,y)\otimes H(x,y)\right] = I(x,y)\otimes\left[\frac{\partial}{\partial x}H(x,y)\right] \qquad \text{A.15}$$

This shows that the derivative of the convolution of an image by a kernel is equal to the convolution of the same image with the derivative of the kernel. In the case of a Gaussian kernel, the derivative can be determined analytically:

$$\frac{\partial}{\partial x}G(x,y) = \frac{-x}{\sigma^2}G(x,y) \qquad \text{A.16}$$

Using convolution with a derivative of Gaussian kernel, it is therefore possible to estimate the gradient of an image.

C. Design and operation of the optimisation framework

The framework was developed early on during the project as a tool to facilitate the screening of very large parameter spaces (order of magnitude ~10^6 combinations), which was required to evaluate various algorithms that were to be developed (or re-implemented from the literature) during the course of the project.

The underlying concept was based on the fact that image processing can be, in most cases, broken down to a series of operations. These operations can include filters (e.g. Gaussian blur), feature description (e.g. connected-component analysis) or the comparison of an image to a manually annotated ground truth. The concept of operation was thus defined loosely as a step that involved some kind of computation based on images or numerical data. Optionally, parameters could be passed to operations and the output was either in the form of images or data (Figure 6-2.A). The implementation was done in MATLAB using handle objects, which are akin to objects passed by reference in other programming languages such as C++.

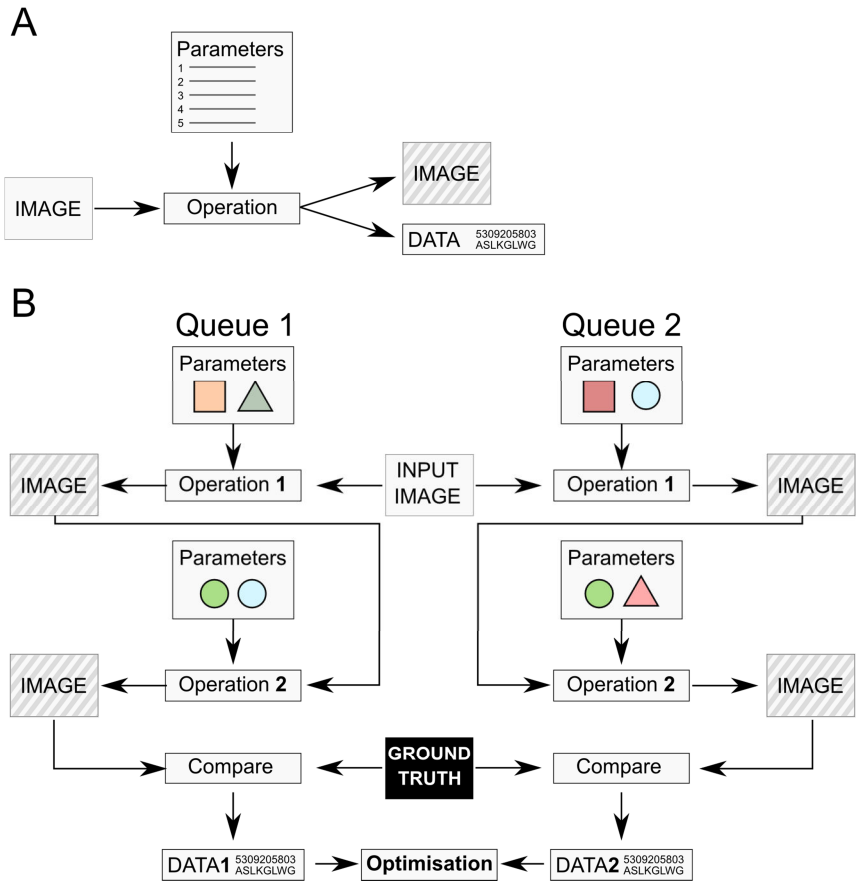


Figure 6-2 Schematic detailing the design and operation of the optimisation framework.

The two main object types that are used to link operations together are the following:

- ***Image Object***: images were represented as multi-dimensional arrays (whose dimensions will depend on the image, e.g. grayscale or RGB). These arrays were embedded in image objects, together with various attributes such as name, original path and data type.

- ***Data Object***: those objects were used to handle data other than images. Data is structured as a table with a String array containing the headers and a generic cell array containing rows of variable data types.

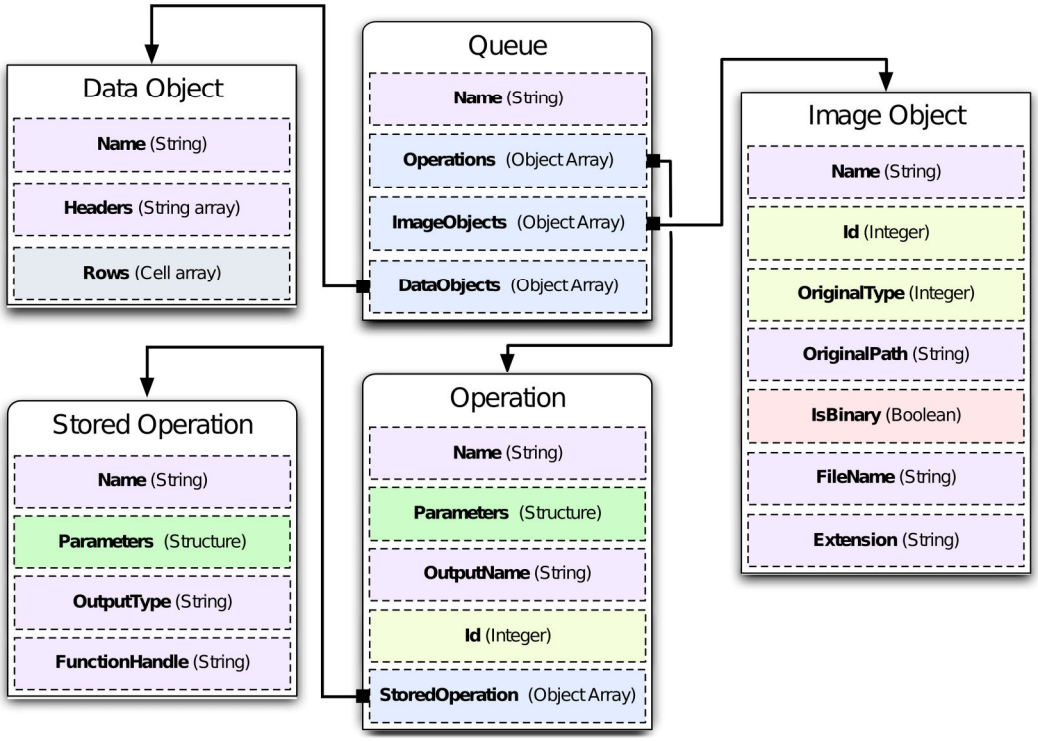


Figure 6-3 Relationship between the main classes of the optimisation framework

Queues are objects that contain the series of operations for a given algorithm. Queues take an input image and produce combination of output images and numerical data, depending on the operations used. During optimisation, one queue is generated per combination of parameter. These queues will share the same chain of operations and be used to process the same input images but will differ solely by the parameter values (Figure 6-2.B). The relationship between the different classes is shown in more details in Figure 6-3.

```
<solution name='Simple local contrast script'>
<!-- The user-generated XML files can containc omments !-->
        <operation name='localContrast' outputName='processed'>
            <parameter name='epsilon' type='range' min='0.01' max='0.2' step='0.01'/>
            <parameter name='image' type='finite' value='mESC'/>
            <parameter name='sigma' type='range' min='0.1' max='2.5' step='0.1'/>
        </operation>
        <operation name='removeSmallObjects' outputName='smallObjectsRemoved'>
                <parameter name='image' type='finite' value='processed'/>
                <parameter name='areaThreshold' type='choices' values='0;100'/>
        </operation>
        <operation name='compareImages' outputName='result'>
                <parameter name='imageRef' type='finite' value='mESC_GT'/>
                <parameter name='imageToTest' type='finite' value='smallObjectsRemoved'/>
        </operation>
</solution>
```

Listing 6-1 Example of a user generated XML script to optimise a simple local contrast thresholding operation

The framework is mainly driven by its graphical user interface (GUI) and user-generated XML (Extensible Markup Language). The latter enabled the creation of complex image processing

algorithm in a few lines by piecing together individual operations. An example of such a script is shown in listing Listing 6-1. In this case, the algorithm is made up of 2 main operations, namely a local contrast threshold followed by a removal of small objects. Both of these operations returned an image while the third one, *compareImages*, returns a numerical array containing the performance metrics computed by comparison with the ground truth image. The framework supported various combinations of image-based and numerical outputs (Table 6-2).

Table 6-2 Return types for framework operations

Type	Description
Image	Images are stored as multi-dimensional arrays whose dimensions depend on their nature (e.g. grayscale, RGB, etc…)
Image Array	Array of image to be used for specific operations such as frame to frame tracking or Z-stacking
Num	Single numeric value
Num array	Array of numerical values

Operations that are accessible by the framework are defined in ***Stored Operation*** objects. These are dynamically loaded upon execution of the framework and can be based either on built-in MATLAB functions or on custom code. Special attributes were used to define parameter values in order to facilitate the determination of optimal parameters (Table 6-3). When parsing the XML file, the framework automatically generated a queue for every possible combination of parameter values.

Table 6-3 Type of attributes for the specification of operation parameters

Value	Description
Finite	The parameter has a single value (defined using the *value* attribute)
Choices	Different values for the parameter are defined using the *values* attribute
Range	A range of values is generated for the parameter (from *min* to *max* with a given *step*)

Once the XML file has been parsed and the objects created, the user can start the parameter space screening, which is done through a grid search. This part of the framework is parallelised using MATLAB's **parfor** syntax and as thus can take advantages of the availability of multiple cores. The results for every queue screened are stored in an easily manipulated format. The main bottleneck with the current version of the framework is mostly memory and storage space. The data of a typical optimisation run can take up around 500MB per input image for 200 000 queues.

As the entirety of the screening data is saved, the optimisation process can be carried offline (i.e. it doesn't have to be done concurrently with the parameter space screening). A simple GUI enabled the user to select the metric to be used for segmentation error computation and to initialise the optimisation using one of the three built-in methods (Figure 6-4), namely single image, whole-set and leave-one-out optimisation (see section 2.2.5 for additional information about these three parameter optimisation methods). Optimisation results were displayed in a

table that could be effortlessly copy/pasted into other software for analysis (e.g. Excel, MATLAB or R).

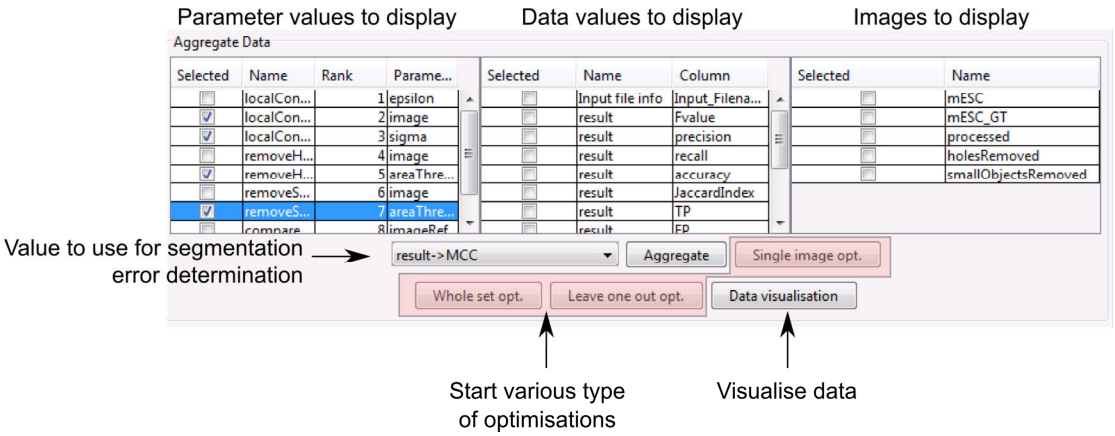
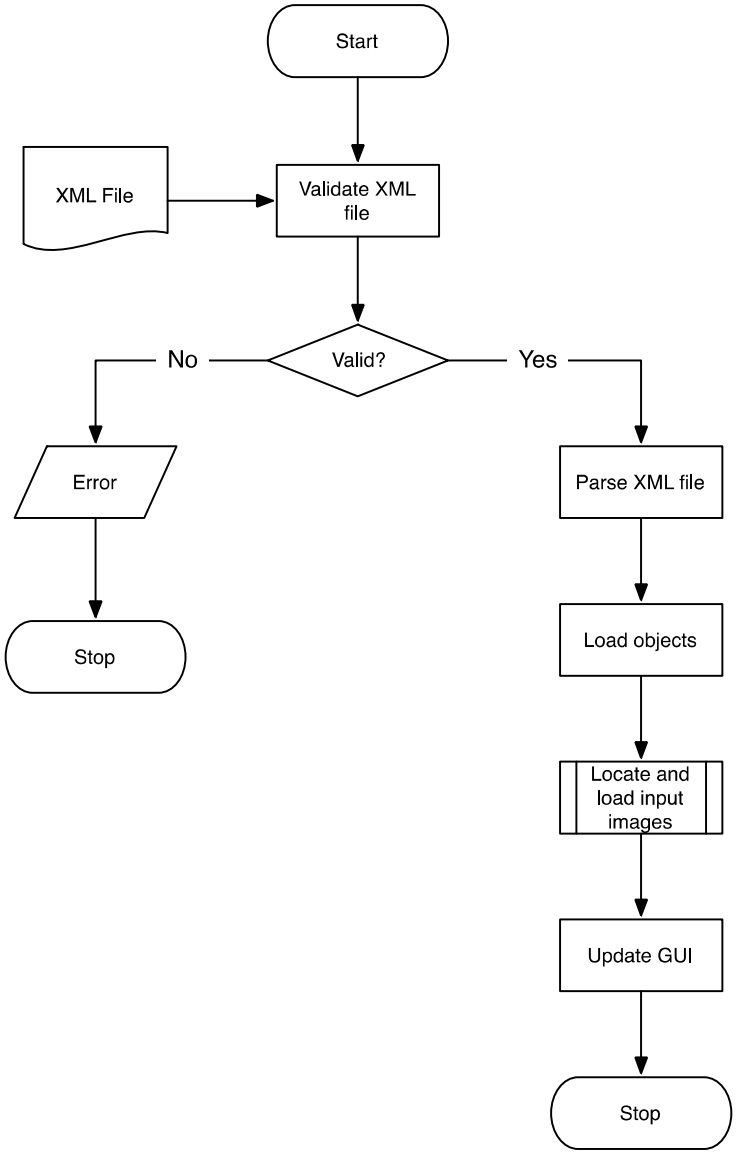


Figure 6-4 Framework GUI for parameter values optimisation.

By combining the ability to easily extend the palette of operations available and the versatility offered by the XML-based scripting, it was trivial to assess the algorithms that were developed and refined as the project progressed.

C.1 Flowcharts describing framework operation

C.1.1 XML parsing and loading



D. Code listings

D.1 Simple Gaussian filter implementation

```
function filteredImage = gaussianFilterCustom(image, sigma)

% Optimal size should be 2.9786*sigma
kernelSize=ceil(2.9786*sigma);
x = -kernelSize:kernelSize;

% Compute the Gaussian kernel
gaussianKernel = exp(- x.^2 / (2*sigma^2) );

% Normalise kernel
gaussianKernel = gaussianKernel .* 1/sum(sum(gaussianKernel));

% Make a copy of the image variable
filteredImage = image;

% Convolve the image in each direction individually

for i=1:2

      % We switch direction of the kernel for the second iteration
      if(i==2)
        gaussianKernel = gaussianKernel';
      end

      filteredImage = …
      imfilter(filteredImage,gaussianKernel,'symmetric','same');
end

end
```

E. Standard deviation computation

For a discrete random variable, the standard deviation for N samples and mean μ is computed as shown in equation E.1.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \qquad \text{E.1}$$

A local standard deviation image *S(x,y)* can be constructed by calculating the standard deviation in $M \times N$ windows centred at each pixel location (M and N both odd integers) as follows:

$$S(x,y) = \sqrt{\frac{1}{M \times N}\sum_{k=x-r_x}^{x+r_x}\sum_{l=y-r_y}^{y+r_y}\left(I(k,l) - \left(\frac{1}{M \times N}\sum_{k=x-r_x}^{x+r_x}\sum_{l=x-r_y}^{y+r_y} I(k,l)\right)\right)^2} \qquad \text{E.2}$$

where $r_x = (M-1)/2$, $r_y = (N-1)/2$ are the 'radius' of the window in the x and y directions, and $\frac{1}{M \times N}\sum_{k=x-r_x}^{x+r_x}\sum_{l=x-r_y}^{y+r_y} I(k,l)$ is the mean intensity value in a window of dimensions $M \times N$ centred at location *(x,y)*.